DAVIDSON ON CONCEPTUAL SCHEMES

by

J.C. Julien Beillard

A thesis submitted in conformity with the requirements for the degree of Ph.D.

Graduate Department of Philosophy

University of Toronto

© Copyright by Julien Beillard (2008)
ABSTRACT: DAVIDSON ON CONCEPTUAL SCHEMES

J.C. Julien Beillard
Doctor of Philosophy
Graduate Department of Philosophy
University of Toronto
2008

In his influential essay “On the Very Idea of a Conceptual Scheme”, Donald Davidson argues that we cannot make sense of conceptual relativism, the doctrine that there could be incommensurably different systems of concepts applicable to a single world. According to Davidson, there is no criterion of identity for language that does not imply or presuppose the possibility that we interpret that language by means of our own language. Given some plausible assumptions, this implies that there is at most one conceptual scheme, one way of interpreting or representing the world. But then the very idea of a conceptual scheme is empty.

The dissertation is an examination of Davidson’s reasoning, and a defence of a different position regarding conceptual relativism. I reject much of Davidson’s argumentation, and his radical (subordinate) conclusion that we would be able, at least in principle, to make sense of any language. Languages that we would be unable to translate or interpret, even in principle, are at least logically possible, in my view. However, this possibility should not be thought to imply or encourage conceptual relativism. In this respect, I think that Davidson and many of his critics have conflated the notion of a difference in conceptual scheme, which requires incommensurability between languages or systems of concepts, and mere conceptual difference.
I argue that a genuinely alternative conceptual scheme would be associated with language unintelligible to us because of its relation to our language. For what is at issue, supposedly, is a conceptual relation: a relation between languages, not a relation between speakers, or their capacities, on the one hand, and languages, on the other. I try to show how some of Davidson’s arguments, suitably modified, can be deployed against the possibility of an alternative scheme, so understood, and provide some additional arguments of my own. My position is thus significantly weaker than Davidson’s: there could not be languages or concepts that we would be unable to understand because they are *incommensurable* with our own.
TABLE OF CONTENTS

Introduction 1

Chapter 1: Quine’s Theory of Conceptual Schemes 24

Chapter 2: A Verificationist Argument? 61

Chapter 3: Organization, Part 1: Extensional Differences 70

Chapter 4: Organization, Part 2: Intensional Differences 102

Chapter 5: Fit, Part 1: Empiricism 120

Chapter 6: Fit, Part 2: Correspondence 157

Chapter 7: Truth and Translation 194

Chapter 8: Error and Interpretation 232

Conclusions 269

Works Cited 307
INTRODUCTION

Donald Davidson contends that languages cannot be intelligibly described as “conceptual schemes”. Essentially, this is to say that languages do not embody or encode radically different ways of interpreting the world. All languages, according to Davidson, can be interpreted by means of our own language: “translatability into a familiar tongue” is “a criterion of language- hood”.\(^1\) Given some natural assumptions, this surprising claim implies that there could not be a conception of the world radically different from our own. Any thinker inhabiting our world must make use of a system of concepts that we could come to understand. So, for Davidson, the very idea of a conceptual scheme or framework that somehow mediates or shapes our experience of the world has no place in a correct account of language and world.

This essay is an examination of Davidson’s argument against the intelligibility of the scheme idea. It is also a defense of a position inspired by Davidson’s discussion, but which differs from his in some important respects. One key difference is that I defend the possibility of a language that we could not translate or interpret. Another is that, unlike Davidson, I think this possibility is consistent with a denial of the possibility of alternative conceptual schemes. I’ll argue that the kind of conceptual inaccessibility required for a difference in scheme is not implied by our inability to interpret certain (possible) languages.

Other differences concern the reasons for rejecting the scheme idea. While I think that Davidson’s main line of argument, with some important qualifications, is basically sound, I find some of his subordinate arguments unconvincing. Where necessary, I’ve replaced these with alternative arguments for similar conclusions. However, some of Davidson’s arguments are directed at views that, as we’ll see, have little bearing on the idea of a conceptual scheme; in these cases no additional argument is needed.

---

\(^1\) Donald Davidson, “On the Very Idea of a Conceptual Scheme”, in *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press), 1984, 186. The essay was originally published in *Proceedings and Addresses of the American Philosophical Association*, 47 (1974): 5-20. Subsequent references to this essay will be to the reprinted version, and will be noted parenthetically in the text as ‘VICS’.
In this introductory chapter, I’ll outline what I take to be Davidson’s treatment of the idea of a conceptual scheme, and the ways in which the position I hope to defend departs from his. The first two sections concern Davidson’s view of the issues. In the third, I describe my own position and summarize the course of my discussion in later chapters.

I

Davidson writes that conceptual schemes are supposed to be “ways of organizing experience” or “systems of categories that give form to the data of sensation”; more generally, they are said to be like “points of view from which individuals, cultures or periods survey the passing scene” (VICS, 183). Of course, it would be absurd to deny that there are points of view, and that people often have different points of view. (Whether it makes sense to say that our concepts or categories organize experience, or that they give form to the data of sensation, is another matter.) What Davidson finds objectionable in talk of conceptual schemes is the additional suggestion that these points of view may differ in some special or radical way, with the result that certain points of view may be mutually inaccessible.

So the idea of a conceptual scheme, in the sense under discussion, involves the idea that there could be a certain kind of difference between ways of conceiving of a single world that would make understanding between thinkers impossible:

There may be no translating from one scheme to another, in which case the beliefs, desires, hopes and bits of knowledge that characterize one person have no true counterparts for the subscriber to another scheme. Reality itself is relative to a scheme: what counts as real in one system may not in another. Even those thinkers who are certain there is only one conceptual scheme are in the sway of the scheme concept: even monotheists have religion. And when someone sets out to describe ‘our conceptual scheme’, his homey task assumes, if we take him literally, that there might be rival systems. (VICS, 183)

Roughly speaking, then, conceptual schemes are supposed to be systems of concepts that may be mutually inaccessible (i.e., mutually “untranslatable”).
These introductory remarks give us a sense of the general kind of view that Davidson intends to explore, and an important hint as to his strategy. Differences in conceptual scheme will be taken to imply the impossibility of translation between the languages that are, presumably, associated with different ways of interpreting or representing the world. Davidson has also stated a key premise in his argument against the intelligibility of the scheme idea: if it makes sense to speak of a “conceptual scheme”, e.g., our scheme, it must also make sense to speak of “rival systems”, of differences in conceptual scheme. If it could be shown that such differences are not intelligible possibilities, it would follow that the idea of a conceptual scheme is empty.

Davidson’s main line of argument against the intelligibility of conceptual relativism (i.e., the view that differences in scheme are possible) stems from an observation made early on in his discussion:

The dominant metaphor of conceptual relativism, that of differing points of view, seems to betray an underlying paradox. Different points of view make sense, but only if there is a common co-ordinate system on which to plot them; yet the existence of a common system belies the claim of dramatic incomparability. (VICS, 184)

According to the conceptual relativist, there can be incommensurably different ways of interpreting or representing “something neutral and common that lies outside all schemes” (VICS, 190). There must be some such entity, for otherwise the idea of a difference in scheme collapses into the idea of thinkers who literally inhabit different worlds, whereas we are concerned here with the idea that the same world could be seen from incommensurably different perspectives. And yet the existence of this common entity would seem to provide a way of making sense of both of these supposedly incommensurable perspectives.

Davidson’s basic objection to talk of conceptual schemes is thus that the idea of a difference in scheme involves two conditions that cannot be met, because they are inconsistent. There must be something common to all schemes, and yet translation must be impossible across languages associated with different schemes. Of course, translation is a particular kind of relation between languages, and it is by no means obvious that we can infer that two languages are inter-translatable from the mere fact that both are related,
somehow or other, to “something neutral and common”. (Or, to put the point another way, we can’t infer that two systems of concepts or ways of conceiving of the world are mutually intelligible from the mere fact that both are related somehow to some shared entity.) So Davidson needs to show that the metaphor of points of view, as it applies to the relations between languages or systems of concepts, on the one hand, and some “neutral and common” entity that these might be thought to interpret or represent, on the other, are sufficient for inter-translatability.

The underlying theme of Davidson’s discussion is that, like the metaphor of points of view, the metaphors of organization and fit are gestures towards real features of conceptual thought or language in relation to the non-conceptual world, features that, on inspection, reveal the contents of thought and language to depend on the world. So the argument against talk of conceptual schemes is not merely destructive, but is also an argument for Davidson’s own view, that mind and world are seamlessly connected. (The phases of this argument will be discussed in somewhat more detail in later sections of this chapter.)

II

Why might one believe that there could be incommensurably different ways of conceiving of the world? One key source of this kind of view lies in Immanuel Kant’s theory of the relation of thought, on the one hand, and non-conceptual experience, on the other. According to Kant, mental life has two distinct components: “a matter for cognition, taken from the senses; and a certain form for ordering this matter, taken from the inner source of pure intuition and thought”.  

Conceptual thought and experience result from the mind’s imposition of structure on something passively received through the senses.

Kant held that there was only one conceptual scheme, but his distinction between conceptual form and sensory content seems to imply the possibility of “rival systems”. For if there is one way of “ordering” the contents of sensory experience, other ways must be at least logically possible. Perhaps some of those other “systems of categories that

---

give form to the data of sensation” (VICS, 183) might result in a conception of the world beyond what we could understand or imagine. Although it is debatable whether this kind of conceptual relativism is a direct logical consequence of Kant’s distinction, it is at least strongly encouraged by reflection on the vast gulf between pre-conceptual sensory experience, as described by Kant, and the kind of conceptual experience and thought of which we are aware. We’ll explore this line of reasoning in more detail in the next chapter, in considering how W.V. Quine’s philosophy might be taken to involve a kind of conceptual relativism.

Although Kant’s philosophy is clearly an instance of the general kind of view under discussion, Davidson is more concerned with more recent accounts of languages or theories as standing to experience or reality in much the same kind of relation as Kant’s system of mental categories. Thomas Kuhn writes, in his highly influential account of the history of science, that scientists trained in different traditions may “occupy incommensurable viewpoints”. In such cases, their (scientific) languages are not inter-translatable: he wonders how scientists in this situation “can even hope to talk together much less to be persuasive”.

Although Kuhn’s position is somewhat unclear in this respect, he also seems to think that reality and truth are relative to a conceptual scheme (or “paradigm”, to use his term). There is some sense, he thinks, in which scientist “after a revolution” in scientific thinking, “work in different worlds” from those of their predecessors, and he is led to doubt whether we can really make sense of the idea of “one full, objective, true account of nature”. Still, he thinks that translation may be possible, if awkward or very difficult, since “the stimuli that impinge on [scientists with different theories] are the same”.

This seems to be a clear example of the kind of view that Davidson describes. Kuhn distinguishes between theories, or the languages used to state and discuss theories, on the hand, and sensory stimuli, on the other. The theories or languages are conceptual schemes, while the shared stimuli, or simply “nature”, provide “something neutral and

---

4 Ibid.
5 Ibid., 135.
6 Ibid., 171.
7 Ibid., 201.
common that lies outside all schemes” (VICS, 190), which can be interpreted in radically different ways. Hence theorists may find themselves in “different worlds”, in some sense, neither being able to make sense of the other’s point of view—even though both inhabit a single world, in some other sense, and their theories purport to describe that same world. So Kuhn’s position seems to involve the “underlying paradox” of conceptual relativism: there is just one “nature”, but perhaps there can be no “full, objective and true account” of nature.

Whereas Kuhn argues for a limited form of incommensurability, said to arise at certain points in the evolution of science, other theorists have held that language generally plays the role of a conceptual scheme. Davidson cites Benjamin Whorf’s claim that “language produces an organization of experience”, that “language first of all is a classification and arrangement of the stream of sensory experience which results in a certain world order”. In much the same spirit, W.V. Quine writes of a distinction between sensory “cues” and the “world view” somehow derived from these:

We cannot strip away the conceptual trappings sentence by sentence and leave a description of the objective world; but we can investigate the world, and man as a part of it, and thus find out what cues he could have of what goes on around him. Subtracting his cues from his world view, we get man’s net contribution as the difference. This difference marks the extent of man’s conceptual sovereignty—the domain within which he can revise theory while saving the data. 

As we’ll see in the next chapter, the “theory”, on Quine’s account, comprises not just scientific theory, in the ordinary sense, but all use of language. Like Whorf, he takes language acquisition and use to involve an imposition of conceptual form on a kind of raw sensory stimulation or experience. This is “man’s net contribution” to his own conception of the world, which is to be distinguished from the world’s contribution. As with Kant’s account of conceptual thought, the world’s contribution is negligible—given the same sensory “cues”, our language or theory could have been very different.

---

III

We’ve briefly noted some influential examples of the general kind of view that Davidson aims to show is not intelligible, and some of the considerations that might lead philosophers to speak of conceptual schemes. In “On the Very Idea of a Conceptual Scheme”, Davidson provides a more detailed account of what he takes to be the sources of the scheme idea, by reference to a pair of related distinctions, the analytic/synthetic distinction and the scheme/content distinction. The first is (or is associated with) “a distinction within language of concept and content” (VICS, 187). Analytic sentences, on this view, are true in virtue of meaning alone, which is to say, in virtue of relations between concepts. Synthetic sentences are supposed to be true in virtue of conceptual facts or facts of meaning together with (non-linguistic, non-conceptual) facts about the world. Hence the “content” of a sentence or group of sentences is the contribution of the world to the truth values (or truth conditions, perhaps) of the synthetic sentences, whatever exactly the contribution may be.

Davidson doubts that an account of differences in conceptual scheme can be based on the analytic/synthetic distinction, since he takes this to imply a “fixed stock of meanings” (VICS, 195) or concepts. This is of no use in attempting to imagine radical conceptual differences. It is unclear what exactly Davidson’s reasoning is in this passage, since a fixed system of concepts is not necessarily a universally shared system of (fixed) concepts. Seemingly, one might hold that schemes differ where one system, associated with certain analytic truths, is incommensurable with another.

Rejection of the analytic/synthetic distinction, however, might seem to provide “a formula for generating distinct conceptual schemes” (VICS, 188). For if there is no such distinction, there is no fixed system of concepts, indifferent to sensory experience or evidence (or the facts, the world). Rather, our concepts, or the meanings of our words, are “contaminated by theory” (VICS, 187), i.e., by the range of sentences we take to be true. On this view, if we come to hold true some range of sentences previously held false, our concepts will also undergo some change.
The change in meaning or concepts cannot be simply that we have come to “view old falsehoods as truths, for a truth is a proposition” (VICS, 188), and what we come to hold true is not the same proposition as what we once held false. So perhaps we should say that there has been a change in “conceptual scheme”, since the whole range of propositions and concepts associated with our language is now different from what it was. This kind of view can be found in the writings of Kuhn, Quine and others; Davidson cites Paul Feyerabend as an example.

The problem with this suggestion is that those operating with different “conceptual schemes”, in the present sense, do not differ in a way that would prevent translation. The purported difference in scheme thus turns out to be trivial. All that is really going on in such cases is that the same words are being used in different ways. Although the reasons for the differences in use or meaning may be interesting or controversial, the difference itself is not. Hence Davidson writes that “instead of living in different worlds, Kuhn’s scientists may, like those who need Webster’s dictionary, be only words apart” (VICS, 189).

Although the line of thought described just now does not seem a promising basis for an account of differences in conceptual scheme, Davidson thinks that talk of differences in scheme is encouraged by a certain response to rejection of the analytic/synthetic distinction. The notion of “content” (or “empirical content”) may be introduced by way of a contrast with the scheme of concepts that make analytic sentences true or false, but one might maintain that there is such a thing as empirical content in the absence of any such contrast: “we can hold, if we want, that all sentences have empirical content”, which “is in turn explained by reference to the facts, the world, experience, sensation, the totality of sensory stimuli” (VICS, 189). Empirical content is then contrasted with an entire language, rather than the purportedly analytic sentences of a language (or the concepts associated with those sentences).

Davidson thus appears to identify the scheme/content distinction with the kind of view that results, in this way, from a rejection of analyticity together with a residual empiricism:

In place of the dualism of the analytic-synthetic, we get the dualism of conceptual scheme and empirical content… I want to urge that this second dualism of scheme
and content, of organizing system and something waiting to be organized, cannot be made intelligible and defensible. It is itself a dogma of empiricism, the third dogma. The third, and perhaps the last, for if we give it up it is not clear that there is anything distinctive left to call empiricism. (VICS, 189).

Certainly the dualism of scheme and content may take the form of an empiricist distinction between sensory intake and conceptual form, or between sensory evidence and theory. But Davidson’s subsequent treatment of this “third dogma” of empiricism seems to involve an attack on a far wider range of philosophical positions. And indeed, we’ve already seen that, in his view, “empirical content” may be identified with “the world” or “the facts” (VICS, 189). Presumably one might hold that languages or schemes represent the world or the facts without endorsing empiricism. These versions of scheme/content dualism, at least, do not seem consistent with Davidson’s account of how the dualism is motivated. In any case, it is clearly wrong to identify the scheme/content distinction with a “dogma” of empiricism.

In later writings Davidson gives a more general account of the distinction. For instance, in “The Myth of the Subjective”, he writes that the contrast between schemes and some shared “content” is involved in the notion of mental representation, and also in the correspondence theory of truth, which he takes to be essentially equivalent:

Content and scheme … come as a pair; we can let them go together… It is good to be rid of representations, and with them the correspondence theory of truth, since it is thinking there are representations that engenders intimations of relativism.10

Here the “content” is not experience or sensory stimulation, but the world or reality. (Davidson later refers to “the dualism of scheme and reality” (VICS, 192). It is clear from his remarks that reality is one candidate for the role of “content”, i.e., something in relation to which schemes might differ.) If language, or, by extension, the mind, represents or corresponds to reality, then perhaps the same reality can be represented in radically different ways. Neither the idea of mental or linguistic representation nor the correspondence theory of truth has any necessary connection to empiricism.

---

Davidson’s attack on the idea of a conceptual scheme, and hence on what he calls “the dualism of scheme and content” (VICS, 189), is thus intended to dismantle not only a central thesis of empiricism, but also a more basic and pervasive theme in modern philosophy. As Davidson says elsewhere, his ultimate target is the “Cartesian” idea that “our perceptions, beliefs … and what we mean by our words … could be just as they are and the world be very different”. As he sees it, the scheme/content distinction lies at the basis of debates about truth, knowledge, meaning, and the relation of the subjective to the objective. While important and interesting in its own right, Davidson’s critique of empiricism is thus a facet of a broader and more radical critique.

IV

In this section, I’ll explain in more detail what I take to be Davidson’s main line of argument against the intelligibility of the scheme idea. As we noted earlier, Davidson approaches the idea of a conceptual scheme indirectly, focusing on the idea of a difference in conceptual scheme. He thinks that talk of a conceptual scheme—e.g., “our conceptual scheme”—implies the possibility of others, ways of thought we would be unable to understand. Perhaps this assumption deserves some defense. Certainly if conceptual schemes are defined as points of view such that “there may be no translating from one scheme to another” (VICS, 183), anyone who holds that we have a conceptual scheme is already committed to the possibility, at least, that there could be some other (i.e., the possibility of a perspective unintelligible to us, given our own perspective).

But Davidson’s position need not depend simply on a stipulation of what the term “conceptual scheme” means. There is a deeper reason for holding that the intelligibility of the scheme idea depends on the intelligibility of the idea of a difference in scheme (i.e., “conceptual relativism”, in Davidson’s sense of the term). Our conceptual scheme (if we have one) is supposed to embody a distinctly subjective element in our experience or conception of the world. This requires that the subjective element could vary to some significant extent independently of how things are in the objective world. So if there could not be other schemes, neither is there any interesting sense in which our conception

---

11 “Myth”, 43.
of the world is subjective: if all (possible) perspectives or systems of concepts were mutually accessible, it would seem that they could, in principle, be integrated into a single, more comprehensive perspective. Our “scheme”, on this view, would simply reflect the world (and our position within it, of course).

The question, then, is whether there could be mutually inaccessible ways of thought. As we’ve seen, Davidson proposes that we focus this question by associating conceptual schemes (if such there be) with languages—more precisely, sets of inter-translatable languages. This provides us with a criterion for differences in conceptual scheme: conceptual schemes differ if and only if the languages associated with which one scheme are not translatable into the languages associated with the other. The impossibility of translation is thus a necessary condition for a difference in conceptual scheme, i.e., for a situation in which one person’s beliefs, perceptions or concepts are inaccessible to another. (Whether it is a sufficient condition is not clear, as we’ll see, but Davidson’s argument requires only the assumption that it is necessary.)

Davidson then pursues the question of whether there could be mutually untranslatable languages by considering the narrower question of whether we can make sense of the idea of a language that we are unable to translate into our language. His aim is to show that all languages are translatable into our language, which would imply that all languages are inter-translatable. The most important phase of his discussion, on the reading I will defend, concerns the question of whether some language might be totally untranslatable into our language. I will ignore for the moment his response to the question of whether a language might be partially untranslatable.

So, in very broad outline, the master argument against the intelligibility of the idea of a conceptual scheme is roughly as follows:

1) The idea of a conceptual scheme is intelligible only if the idea of a difference in conceptual scheme is intelligible.

2) The idea of a difference in conceptual scheme is intelligible only if there could be languages that are not inter-translatable.

3) All languages are translatable into our language.

Therefore,

4) The very idea of a conceptual scheme is not intelligible.
The most interesting and controversial premise is (3), and Davidson provides various subordinate arguments for it. These subordinate arguments, which will be my main focus in what follows, are directed at various attempts to flesh out “dominant metaphor of conceptual relativism, that of differing points of view” (VICS, 184) in a way that would satisfy both of the conditions for a difference in conceptual scheme. Davidson identifies two main groups of metaphors associated with the idea of a difference in scheme, the metaphor of organization, and the metaphor of fit. These metaphors are supposed to give us some sense of how languages could be conceptual schemes. That is, they are supposed to provide us with a way of conceiving of languages in relation to a common world or range of sensory experience, which does not presuppose or imply the possibility of translating these languages, so conceived, into our own language.

Davidson tries to make literal sense of the various metaphors in these two main categories, and then considers whether the relations between language and “content” that seem to be indicated by these metaphors can be understood without introducing the possibility of translation across putatively divergent schemes. Can we describe languages as organizing or fitting some common entity in a way that is independent of the possibility of translation?

First, Davidson considers the metaphor of conceptual or linguistic organization, whether of experience or reality. He argues that the concept of organization, applied to language, seems to concern “the referential apparatus of language” (VICS, 193). In particular, what seems to be at issue is the notion of predication, i.e., of categorizing or classifying some range of entities. This is intelligible, but Davidson argues that any two languages that serve to categorize or classify a shared range of objects must be inter-translatable. The shared ontology of the two languages guarantees the possibility, at least in principle, of discovering the extensions of terms associated with a purportedly alien scheme and mapping them on to our own language. For instance, if some alien language categorizes people by their distances from Saturn instead of by income or occupation, say, this would certainly count as a different way of “organizing” people. But we could discover and understand how the alien language differs in this way from our own.

The situation is essentially the same whether we take these objects to be constituents of some form of “experience” or simply the extra-linguistic world. The only
difference is that the first possibility suffers from the additional problem that, in Davidson’s view, there could not be a language that “organizes” only experiences. So this suggestion either collapses into the view that languages organize reality (together with whatever portions of reality we may call “experience”), or it is untenable for independent reasons.

Davidson then turns to the idea that languages or schemes “fit” some entity, whether experience or reality. He takes this metaphor to concern the role of complete sentences rather than the sub-sentential apparatus of reference. One might hold that sentences are warranted by the empirical content, e.g., by sensory experience or evidence. This claim also is intelligible—or may be intelligible, on a certain interpretation—but Davidson argues that it can be reduced to the idea that experience or evidence “makes sentences and theories true” (VICS, 94). But the claim that sentences or theories are made true by evidence or experience, he thinks, reduces to the bare claim that those sentences or theories are true: “the notion of fitting the totality of experience, like the notion of fitting the facts, or of being true to the facts, adds nothing intelligible to the simple concept of being true” (VICS, 194).

The question of whether there could be an alternative scheme, so understood, thus becomes the question of whether there could be a theory or group of sentences that is both true and not translatable into our language. This is impossible, according to Davidson, because the concepts of truth and translation are essentially inter-dependent. Tarski’s Convention T determines the extension of the concept of truth for any language, and does so “by making essential use of the concept of translation” (VICS, 194): the meta-language sentences that state truth conditions for the object language (or theory, i.e., group of sentences) are, in effect, translations of sentences in the first language. Davidson concludes that, because “Convention T embodies our best intuition as to how the concept of truth is used” (VICS, 195), the present account of conceptual schemes cannot be made intelligible. In conceiving of something as a true theory, i.e., a true group of sentences, we are committed to the possibility of translating that theory into our language.

The argument from Convention T concludes what I take to be the most important phase of Davidson’s discussion. (Arguably, that argument is the ultimate basis of
Davidson’s attack on the scheme idea.) To summarize, Davidson holds that the scheme/content distinction arises from a certain response to the rejection of the analytic/synthetic distinction: schemes are to be associated with languages, and contrasted with the “empirical content” inherited from the analytic/synthetic distinction. He then argues against the intelligibility of various accounts of the relation between schemes and content. His view is that, insofar as we can understand languages as standing in the relations under discussion, we are forced to conclude that languages so related to a shared “content” must be translatable into our language, and hence inter-translatable. This leads to the further conclusion that talk of conceptual schemes has no content or purpose.

The remainder of Davidson’s discussion addresses the idea of a partial difference in conceptual scheme: a language that is partially untranslatable into our language. In my view, this is a less important part of Davidson’s discussion, although, as we’ll see, it introduces some important themes in his philosophy from which additional arguments against conceptual relativism can be drawn. Davidson’s basic position regarding the idea of a partial difference in scheme is that “the underlying methodology of interpretation” (VICS, 197) rules out the possibility of discovering a partially untranslatable language. Interpretation, for Davidson, requires that we maximize the truth and rationality of the beliefs we ascribe to others. Consequently, anyone whose language we can partially interpret must also be a speaker whose concepts and beliefs we take to be largely the same as our own. This does not mean that we cannot discover disagreements or confusions, of course, but it does rule out the possibility of radical differences.

V

The position I’ll defend in this essay differs in some important respects from Davidson’s position, or at least from what is commonly taken to be Davidson’s position. I’ve already indicated the most important differences: I think that there could be languages that we would be unable to translate or interpret, but I don’t see this as equivalent to the possibility of alternative conceptual schemes. However, I should emphasize that I also disagree in some important respects with certain common readings
of Davidson’s argument. The argument is stronger and more interesting, in my view, than it is often taken to be. In this section, I’ll expand on these differences between my reading of Davidson’s argument and certain others, and on the differences between what I take to be his position and my own.

One fairly common objection to Davidson’s argument, as we’ll see, is that it depends on some form of verificationism, i.e., on some premise or assumption to the effect that our ability to recognize something as a language, by translating or interpreting it, is constitutive of its being a language. On the reading I favor, this is not a premise or assumption of the argument, but a conclusion, as Davidson says. (That “translatability into a familiar tongue” is “a criterion of languagehood” is supposed to “emerge as the conclusion of an argument” (VICS, 186), according to Davidson.) In effect, Davidson is arguing for a kind of local verificationism, about language and the related concepts of meaning, belief and thought, but the argument does not depend on any question begging assumptions. Nor does it depend on the assumption that meaning or truth, in general, are somehow reducible to conditions of verifiability, warranted assertion, etc.

A second difference, related to the first, is that I do not attach much importance to Davidson’s appeal to inter-translatability as a criterion for identity of conceptual scheme. I think that Davidson intends the notion of translation to be understood very broadly, perhaps so broadly that the term “translation” is misleading. So Davidson’s view is presumably that interpretation or understanding across languages is the criterion for identity of conceptual scheme. Understanding might come by way of translation, but it need not. Instead, one might simply learn a new language, for example. (As we’ll see, some commentators have made much of the fact that two languages, or two fragments of a single language, may not be inter-translatable—not, at least, in a certain sense of the term. I think this is a red herring. It is trivially true that such languages are possible, and tells us nothing about the possibility of understanding.)

Finally, the reading of the argument that I defend does not imply, in particular, that we could come to understand an alien language or system of concepts only by discovering it to be essentially the same as our own. Perhaps this point deserves clarification. What I deny is the claim that if one can understand a view of the world other than one’s own, it is necessary that that view of the world consists, at least for the
most part, in beliefs and concepts that one already possesses, or, at least, that one takes that other view to consist in those kinds of beliefs and concepts. I think all that is required for understanding is that the other view be taken to consist in beliefs and concepts that can be integrated or added to the interpreter’s existing system, with the result that both systems can be seen as parts of a more comprehensive system.

Suppose, for instance, that some isolated tribe of hunter-gatherers are introduced to modern civilization, and come, over time, to understand our conception of the world. They learn about countless things that, from their perspective, will be extremely novel and surprising, and hard to even understand: computers, stock markets, natural selection, modernist poetry, and so on. Perhaps there is a sense in which we could say that they have discovered our view of the world to be essentially the same as their (earlier) view of the world. For there will presumably be certain very basic concepts and beliefs common to our view and theirs: all of us have the concept of a person, the concepts of action, belief and intention, and other concepts, and all of us believe that there exist trees and rocks, that we need water and sleep, and so on. But the fact that their view and ours can both be framed in terms of these very general and basic concepts is consistent with the fact that vast expanses of our world view are, from their perspective, completely new and alien—and so, in this sense, very different from theirs.

Davidson’s position regarding this kind of situation is not clear. It is an obvious empirical fact that differences of this kind exist, so presumably Davidson does not mean to deny that they do, let alone that they could exist. But then it is hard to know what to make of his claim that “given the underlying methodology of interpretation, we could not be in a position to judge that others had concepts or beliefs radically different from our own” (VICS, 197). If a “radically different” system of concepts and beliefs were one that we find it impossible to understand or interpret, this claim would be tautologically true: we could not be in a position to interpret what is not interpretable. But if it is hard to believe that Davidson is denying obvious empirical facts, it is equally hard to believe that he means to put forward a tautology as a thesis about mind and world with important philosophical consequences.

So it would seem that Davidson’s notion of a “radical” difference in concepts or beliefs must involve a criterion weaker than the impossibility of interpretation but
stronger than the impressive degree of difference we find in the actual world. Somewhere between the actual and tautological cases, there is some degree of difference that Davidson takes to be “radical” and, he thinks, not possible. I can see no reason to believe this, and no principled basis for holding that some particular degree of difference is not possible. Given what we know of the actual world, we can surely imagine greater differences of the same kind.

We can even imagine, I think, that certain speakers might have a system of beliefs and concepts that does not overlap in any substantive way with our own. Perhaps simply in virtue of having a language they would have certain concepts that we also have, such as the concepts of truth and belief. But they might communicate only about aspects of the world that are as utterly alien to us as electricity or psychoanalysis would be to the hunter-gatherers. The thesis I want to defend, which can be drawn from a certain reading of Davidson’s discussion, is merely that where two world views differ in this way, they can be integrated into some more comprehensive view, which need not be identified with either of the first two.

From the perspective of the third view, language or system of concepts, the first two can be seen as more or less complementary accounts of a single world, or as more or less conflicting accounts. I don’t think that it is necessary that this third world view be held by all possible thinkers or speakers, or that this view be held by any actual speakers. But the mere fact that such a view exists, or could exist, shows that there is no difference in conceptual scheme in such cases.

I now turn to some differences between my own position and what I take to be Davidson’s position, as opposed to what other commentators have taken it to be. The most important difference concerns the modality of the thesis that there could not be a language that is uninterpretable by means of our language. Does this mean that we would be able to interpret or understand any logically possible language? Earlier I said that I disagree with Davidson on this point, in that he holds that we would be able to understand any possible language, while I deny this. In fact, it seems to me that there is a reading of Davidson’s argument, or at least of certain key phases of his argument that supports my view of this matter. However, it is also fairly clear that Davidson himself does not, or does not consistently, read his own argument in this way. So the line of
argument I defend is one that is inspired by Davidson’s argument, on a certain reading, but it is at odds with his understanding of the way in which the argument works, and what it establishes.

In my view, what is really at stake as concerns the idea of a difference in conceptual scheme is a relation between concepts, hence between the languages associated with putatively divergent schemes. This is what “incommensurability” should be taken to mean, I think: it means that there is no common measure of the concepts or languages in question, not merely that certain speakers are unable to discover a common measure. So if schemes differ, and two speakers are therefore unable to understand or interpret each other’s utterances, this must be due to the special way in which their respective concepts or languages are related (i.e., the relation of “incommensurability”).

But this is not the only possible situation—if it is a possible situation at all—in which one speaker is unable to understand another, interpret or translate the other’s language. There could be languages, for example, that are learnable only by reference to sensory modalities or properties that we lack. Therefore, I think that some such logically possible languages, at least, are languages we would be unable to interpret. But in this kind of case, the languages themselves may be commensurable. That is to say, it would be consistent to suppose that some more perceptive or intelligent being would be able to interpret both our language and that of the aliens in his language.

Such an interpreter would find, we might suppose, that the aliens know certain truths that are not knowable to us (and perhaps that we know certain truths unknowable to them). But, as I’ve said, the mere possibility of a third perspective or language that could integrate both of the first two as elements of a more comprehensive view of the world implies that the first two are not incommensurable. There is a common measure, in this case, but we may not know it or even be capable of knowing it. So the mere fact that we are unable to translate, interpret or understand a language does not imply that that language is incommensurable with our own, i.e., that it is associated with an alternative conceptual scheme.

My position, then, is that any possible language is such that it can be interpreted by means of our language: we might not be able to carry out the interpretation, but some (possible) interpreter could do so. This is a weaker reading of the claim that there could
not be a language that is uninterpretable by means of our language. Again, because the real issue here concerns the relations between languages or systems of concepts, in my view, a difference in scheme would be a situation in which one speaker’s language prevents him from understanding another language.

Davidson’s general strategy against talk of conceptual schemes can be seen to imply or at least suggest this important qualification—although, as I’ve said, his position is neither entirely clear nor consistent. In trying to develop a reading of Davidson’s argument that I find persuasive, I have emphasized those features of his discussion that can be used to defend my view that the idea of a difference in conceptual scheme requires not just the impossibility of interpretation, but an impossibility resulting from a relation between concepts or languages. Yet it seems fairly clear that Davidson denies the possibility of a language that we could not interpret, as we’ve seen.

However, I should emphasize that Davidson’s view, on my reading, does allow for the possibility that we are unable at present, or for a very long time, to interpret a language. For instance, we might at present lack the knowledge to interpret a language associated with some theory far more sophisticated than any we presently have. As we’ll see, some commentators take Davidson to deny even this possibility; of course, such a position would be simply absurd. What Davidson claims is rather that, given sufficient time and effort and ingenuity, we could come to understand any language, at least in principle.

Finally, I am not convinced by Davidson’s alternative to the dualism of scheme and content, a view that treats the concepts of truth, meaning, belief and interpretation as equally basic and inter-dependent concepts. I think that some modest form of realism is tenable without any form of scheme/content dualism, and is a more attractive alternative. On this view, which seems to be at odds with Davidson’s, the concepts of truth, meaning, belief and interpretation are intelligible because we take ourselves to inhabit a single world that is largely independent of our beliefs and concepts. This anodyne form of realism may be essentially linked with the concept of truth and, by way of truth, the concepts of belief and meaning; but these concepts are not dependent on the concept of interpretation.
VI

The course of my discussion will be as follows. In Chapter 1, I explore some themes in W.V. Quine’s philosophy that might be taken to encourage conceptual relativism. This will serve to clarify what is at stake, but it is also of interest because of Quine’s significant influence on Davidson. In considering Quine’s views, we will also introduce some basic themes in Davidson’s writings on mind, meaning and language. Chapter 2 is a fairly brief discussion of the verificationist reading of Davidson’s argument, and of my reasons for rejecting this line of argument. However, the issue of verificationism will also be considered from time to time in later chapters, since it is a pervasive issue in discussions of Davidson’s argument.

In Chapter 3, I will begin a detailed examination of what I take to be the major subordinate arguments. The first is an attack on the idea that conceptual schemes can be understood as systems of organizing principles. I explain Davidson’s objection to a certain interpretation of this claim, which concerns the extensional features of language, and conclude that the objection is sound. In the course of the discussion I argue against some important misunderstandings of Davidson’s strategy against this account of differences in conceptual scheme, and against the suggestion that mere conceptual diversity could be considered sufficient for a difference in scheme. Then in Chapter 4 I address a possibility that Davidson ignores, the possibility that the notion of conceptual organization concerns the intensional features of language. This may seem to be a promising basis for an account of differences in conceptual scheme, and perhaps a more natural construal of the notion of organization. I argue that this suggestion fares no better than the extensional construal that Davidson attacks, although for different reasons.

Chapter 5 concerns Davidson’s argument against a different formulation of the scheme/content distinction, which draws on some form of knowledge empiricism. On this view, conceptual schemes do not organize experience or reality, but instead “fit” some entity. I argue that Davidson’s treatment of this idea is confused; it depends on a chain of conceptual assimilations that cannot be fairly attributed to proponents of the scheme idea, and which have little intrinsic plausibility. I focus in this chapter on the first phase of the argument, which purports to reduce the notion of ideal empirical
warrant to the notion of truth. While this argument is unconvincing, for various reasons, a better and simpler argument against accounts of conceptual schemes based in knowledge empiricism can be drawn from Davidson’s critique of the empiricist view that non-propositional entities such as sensations or sense data could serve as a kind of evidence for beliefs or theories.

I turn in Chapter 6 to an examination of the second phase of Davidson’s argument against accounts of conceptual scheme based in knowledge empiricism. Although the argument is unsound for reasons considered in Chapter 5, it is worth considering further so as to better understand Davidson’s reasoning, and also because the second phase of the argument can be treated as an independent argument against the very idea of a conceptual scheme. Here the issue is the relation between the scheme idea, on the one hand, and the notions of conceptual representation and correspondence to fact, on the other. I explain and evaluate the slingshot argument, a collapsing argument often attributed to Frege, which Davidson takes to have devastating consequences for the notions of correspondence, representation and conceptual scheme. Although I agree with Davidson’s skepticism regarding traditional theories of correspondence, I dispute his account of the relation between these notions. Whereas Davidson thinks that the idea of correspondence to fact or reality encourages conceptual relativism, I hold that a correspondence theory of truth would provide a powerful argument against the idea of a difference in conceptual scheme.

Chapter 7 addresses an argument that Davidson treats as the third phase of his larger argument against accounts of differences in conceptual scheme based on knowledge empiricism. Conceptual schemes were to be associated with theories borne out by the same sensory evidence. Davidson claims that this suggestion reduces to the idea that they correspond to the evidence, or to the facts, and that this, in turn, reduces to the bare claim that they are true. He then argues that the concept of truth is essentially dependent on that of translation, by appeal to the role of translation in Tarski’s Convention T. Since the two attempted conceptual reductions that lead to this phase in the argument fail, the argument is, of course, unsound.

However, the third phase of the argument, like the second, can be detached from the larger argument to which it is supposed to contribute. We can consider instead the
suggestion that alternative conceptual schemes are true theories that are not translatable, or, more precisely, interpretable, by means of our language or system of concepts. This deserves consideration since at least some of the beliefs or theories associated with a purportedly different conceptual scheme must presumably be true, or true “relative to” that alien language, theory or view of the world.

I consider a number of possible readings of the argument from Convention T, and conclude that none is sound. The most interesting reading of the argument requires the assumption that all linguistic understanding is interpretive, a central tenet of Davidson’s philosophy. I argue that this is false, and that the concept of truth, unlike that of truth in a language, is not essentially dependent on that of translation or interpretation. I also point out a difference between ways of understanding what is expressed by instances of Convention T that Davidson appears to ignore, and suggest an account of the relation between the Tarskian truth concepts and the more general concept of truth that, in my view, accounts for the facts better than Davidson’s. I then suggest a different, simpler argument against the intelligibility of the account of differences in conceptual scheme under discussion.

Chapter 8 concerns two different but closely related matters. First, I evaluate Davidson’s treatment of the idea of a partial or local difference in conceptual scheme. It is unclear from Davidson’s discussion what exactly he takes to be involved in this idea. After some rejecting some initial possibilities, I consider the suggestion that an otherwise interpretable language might contain a fragment that is uninterpretable. Davidson’s argument against this suggestion seems to be based in the holistic nature of interpretation. In my view, the argument is unsound.

I then turn to a seemingly unrelated issue, the question of whether our view of the world could be radically mistaken. I explain how Davidson’s view of the methodology of interpretation is supposed to rule out this possibility, on the basis of the holistic nature of belief and meaning discussed earlier. After some discussion of his argument from the possibility of an omniscient interpreter, I suggest that the case against the possibility of massive error hinges ultimately on a moderate form of content externalism. Together with moderate holism, moderate content externalism rules out the possibility of massive error (as well as a more radical or total holism).
Although Davidson defends this form of externalism on the basis of the methodology of interpretation, I argue that this is another reversal of conceptual priorities: it is precisely because language or concept acquisition cannot be a kind of interpretation that content must be externally determined to some significant extent. Rather the methodology of interpretation, the impossibility of massive error and the impossibility of divergent conceptual schemes are all consequences of this moderate externalism. To the extent that moderate holism and externalism can be held to be intrinsic to our understanding of what it is to have beliefs and concepts and use language, this is a fairly powerful argument, in particular, against the possibility of a difference in conceptual scheme.
CHAPTER 1
QUINE’S THEORY OF CONCEPTUAL SCHEMES

In this chapter, I’ll try to show how Quine explains and defends a version of the scheme/content distinction. There are certain basic themes in Quine’s writings that seem to invite conceptual relativism: the radical under-determination of theory by evidence, and the indeterminacy of translation. Sometimes Quine takes these doctrines to show that there are, or at least that there could be, different conceptual schemes. More often, however, he concludes that the idea of an alternative to our own conceptual scheme is empty or unintelligible, for reasons that closely anticipate Davidson’s critique. This ambivalence is symptomatic of a serious and perhaps irresolvable tension between Quine’s empiricist theory of knowledge and his empiricist theory of meaning—and, ultimately, between empiricism and naturalism. As we’ll see, some of Davidson’s objections to the scheme idea touch on similar issues.

I

Quine has no official theory of conceptual schemes, or even of concepts. Concepts, like meanings or propositions, can’t be individuated or characterized in the behavioristic and extensional terms that Quine favors. In a brief reply to Davidson’s charge that he is committed to an invidious distinction between conceptual schemes and empirical content, Quine explains that his use of the phrase “conceptual scheme” is a figure of speech, playing no technical role in his system.12 As we’ll see, this line of defense is not very plausible. The idea of a difference in conceptual scheme is central to Quine’s philosophy, and the source of much tension and perhaps outright inconsistency between Quinean doctrines.

---

Quine’s main interest is the relation between science and sensory evidence. Science is a system of sentences, a set of inscriptions and utterances, together perhaps with the complex behaviors and dispositions that figure in language use. The system as a whole is keyed indirectly and holistically to the activation of nerve endings. If we want to inquire more closely into the relation between this linguistic system and its sensory basis, we will find that this is in large measure the province of “neurology, psychology, psycholinguistics, genetics [and] history”. There is, however, a purely philosophical question concerning the relation of evidential support, “and its essentials can be schematized by means of little more than logical analysis”.

So Quine holds that the scheme/content distinction is more or less the distinction between theory and evidence. This distinction, in turn, is basically the distinction between (a) certain kinds of behavior and dispositions to behavior, and (b) the neurophysiological causes of those behaviors and dispositions that are most directly connected to the world outside our bodies. It may be difficult to formulate the distinction precisely; but it is surely not, as Davidson holds, unintelligible:

The third purported dogma [of empiricism, i.e., the scheme/content distinction] … has both a descriptive and a normative aspect, and in neither aspect do I think of it as a dogma. It is what makes scientific method partly empirical rather than solely a quest for internal coherence. It has indeed wanted some tidying up, and has had it.

The descriptive aspect of the “dogma” is that our knowledge of the external world is ultimately based in the activation of nerve endings. This Quine takes to be a discovery of natural science; studies in physiology and empirical psychology have vindicated an empiricist orientation in epistemology. The normative aspect is a consequence of the descriptive. Given what we know about knowledge, we are rightly skeptical of theories unconstrained by sensory stimulation—numerological or astrological theories, for example. Neither the empiricist finding nor the empiricist norm is an uncritical “dogma”. Both are fallible, scientific hypotheses, subject to empirical refutation. Thus Quine allows that it is possible, although of course absurdly improbable, that we might come

---

14 Ibid., 2.
15 “Third Dogma”, 41.
across convincing evidence of clairvoyance or other “extra-sensory” perception. In that case, we would have reason to reject empiricism.

There is, then, a kind of circularity in Quine’s defense of empiricism. Empiricism is supported by science, but the authority of science derives from the empiricist norm: science is uniquely sensitive to empirical constraints. Thus it seems that Quine’s argument reduces to the claim that if empiricism is warranted, so is natural science, and vice versa, but he offers us no independent reason to think that either is warranted. Quine’s answer to this kind of objection is what he calls “naturalism”, a repudiation of “the Cartesian dream of a foundation for scientific certainty firmer than scientific method itself”. 16 We can only ask and answer questions against the background of some theory or other, and if we are serious the theory can only be the best that we have at present, our current natural science. The alternatives are merely other, less warranted theories. Therefore, it simply makes no sense to suppose that we could set aside all of our best warranted beliefs and retain the ability to judge what is or is not warranted.

So, on Quine’s view, it makes no sense to suppose that we could assess the methods and findings of science without reference to those very methods and findings. Circularity of a certain kind is thus inevitable, but perhaps innocuous. Consider by analogy the skeptical question of why we should think logically. Any persuasive argument for the authority of logic will, by its very nature, presuppose the authority of logic—to whatever extent it is persuasive, it will at least appear to be logically valid. Some might take this to suggest that there is ultimately no reason to think logically. A more plausible conclusion might be that logic is constitutive of thought. The authority of logic can’t be defended on non-logical grounds because the authority of logic is built into the practice, and thus the very concept, of justification.

Similarly, Quine holds that our more general epistemic criteria and concepts are rooted in a particular theory of the world. There is no way to justify that theory as a whole, or to call it into question, except by appeal to the theory itself. This would be to demand, as it were, a reason to be reasonable. This is also, in essence, Quine’s answer to radical skepticism:

---

16 Pursuit of Truth, 19.
We cannot significantly question the reality of the external world, or deny that there is evidence of external objects in the testimony of our senses; for, to do so is simply to dissociate the terms ‘reality’ and ‘evidence’ from the very applications which originally did most to invest those terms with whatever intelligibility they may have for us.\(^{17}\)

And yet he also holds that skepticism is a false empirical hypothesis. For the evidence that leads to skeptical doubts is empirical evidence, consisting of optical illusions, hallucinations and the like. Thus we are free to appeal to science in refuting skepticism. Again, the normative and the factual blend together. On the one hand, we have good empirical evidence that the external world is real and known to us by “the testimony of our senses”. On the other hand, these empirical beliefs are constitutive of the very concepts of reality, evidence or knowledge. The striking conclusion appears to be that while it is a contingent, empirical truth that our theory is warranted by sensory evidence, it would be *incoherent* to imagine that it isn’t. There is certainly a tension here, one which we’ll see is central to Quine’s philosophy.

Quine’s project seems essentially intelligible and defensible. It is, as he often says, merely the scientific study of the nature and methods of science. Traditionalists are free to pursue “the Cartesian dream” of an indefeasible kind of evidence or reason against which we could measure our beliefs. But surely Quine is not to be faulted for concluding that this is futile. Naturalized epistemology does not offer an account of knowledge that will satisfy traditionalists or refute radical skepticism, but it may offer an account of knowledge nonetheless. This account has no special status, higher than that of any natural science; it is only as defensible as physics, biology or history—which is to say, highly defensible.

This is not a complete characterization of Quine’s naturalistic empiricism, however. The destructive thesis that underwrites Quine’s system, as we’ve seen just now, is that there is no higher theory or methodology to which natural science is answerable. I have suggested that this is a defensible view, at least at first blush. Quine’s constructive program, however, goes far beyond what would normally be counted as

natural science. In large measure, it is an attempt to reinterpret traditional questions and problems of philosophy in what Quine takes to be a properly scientific fashion. It is unclear, to say the least, whether this is a straightforwardly scientific project, as Quine takes it to be.

Quine does not simply ignore traditional philosophical concepts, questions and problems. Rather, he aims to *incorporate* them into natural science by recasting them in terms he deems properly scientific—to graft philosophy on to empirical science, to whatever extent possible. It is not obvious that this can be done, or that the project even makes sense. This is the source of the tension noted earlier. In defending empiricism, or answering skepticism, Quine seems to conflate what we would normally consider empirical and conceptual claims, in a puzzling manner. If skepticism, for example, is simply unintelligible, resting on a misuse of words or concepts, how could it be that it is also a trivially false scientific hypothesis?

Quine’s allusion to Kantian philosophy (“the transcendental abyss”) is illuminating. On the one hand, Kant set himself a very similar task: the task of determining the limits of thought, of salvaging or dissolving metaphysical pseudo-problems by recasting them in a new form. On the other hand, Kant’s strategy presupposes a distinction that Quine vehemently rejects, between the empirical and the transcendental. Thus Kant holds that spatio-temporal objects are empirically real but transcendentally ideal: they are real and independent of any mind from the perspective of ordinary experience and science, but ideal and somehow dependent on the activity of the knowing subject from another perspective. Quine’s naturalism is the denial that there is any such transcendental perspective; there is only the empirical perspective of natural science.

So, like Kant, Quine claims to be salvaging metaphysics by recasting it within what he takes to be the limits of thought. Unlike Kant, he identifies the limits of thought with the limits of empirical science:

We must speak from within a theory, albeit any of various. Transcendental argument, or what purports to be first philosophy, tends generally to take on this status of immanent epistemology insofar as I succeed in making sense of it. What
evaporates is the transcendental question of the reality of the external world—the question whether or in how far our science measures up to the Ding an sich.\(^\text{18}\) This is not quite right, as we’ll see. The “transcendental question” does not entirely disappear; Quine thinks that it is not utterly unintelligible. Something like Kant’s distinction between the immanent, empirical perspective and the transcendental perspective can be reformulated as “immanent” epistemology. We can measure the extent to which our theory of the world measures up to the world described by that same theory. More precisely, we can study language acquisition and theory construction by means of the theory being studied.

Quine suggests that this affords us a quasi-transcendental perspective, insofar as science itself provides a measure of something like Kant’s Ding an sich—the basic features of objective reality are described by physics. As he explains, the very concept of a “fact of the matter” is “to be taken naturalistically within our scientific theory of the world”: it is the notion of “distributions of states and relations over elementary particles”.\(^\text{19}\) Since physics is quasi-transcendental, we can in some sense measure our theory of the world against the world as it is in itself. For example, we can see that talk of propositions, properties and mental states is not to be taken literally; it reflects no physical fact of the matter. In this way, Quine aims to distinguish between the genuinely objective and the merely subjective features of our language, theory of the world or conceptual scheme.

This line of thought is deeply puzzling. The assertion that “we must speak from within a theory” seems to be made from a perspective transcending that theory, or any. For how could it be intelligible to say that we see things from a certain perspective, except by contrast with some conception of what the world is like apart from that perspective? If all of our thoughts and perceptions are “immanent” to our current theory, there is presumably no way for us to know that they are, and no point in saying so. The claim that they are “immanent” could only be understood, in that case, within the same conceptual or epistemic limits under discussion. Likewise, to recognize that we may speak from within “any of various” theories is presumably to consider these different

\(^{18}\) W.V. Quine, “Things and Their Place in Theories”, in *Theories and Things*, 22.

\(^{19}\) Ibid., 23.
theories in relation to some common object that they represent in their different ways. But this, again, seems to imply that we have some access to the world that transcends our own theory, or indeed any theory. Or so one might argue, at least.

Indeed the situation is perhaps worse than I have suggested so far. Quine’s slogan that “we must speak from within a theory” is not simply the innocuous and uninteresting claim that we can only ask and answer questions against the background of some particular set of beliefs. His discussion of our “immanent” perspective suggests, in addition, that we are somehow confined by our beliefs—that our view of things is necessarily subjective, in some important and disquieting sense. But he also holds that truth is “immanent”, a relation between theories and theoretical posits:

Where it makes sense to apply ‘true’ is to a sentence couched in terms of a given theory and seen from within that theory, complete with its posited reality… To say that the statement ‘Brutus killed Caesar’ is true, or that ‘The atomic weight of sodium is 23’ is true, in effect simply to say that Brutus killed Caesar, or that the atomic weight of sodium is 23.²⁰ Brutus, Caesar, sodium and atomic weight are all elements of the “posited reality” of our theory. Thus to say that “Brutus killed Caesar” is true is just to say that Brutus killed Caesar; but that is just to say how things are in the “posited reality” of our theory. For the sentences of any theory “are about posited entities, are significant only in relation to a surrounding body of theory”.²¹ Of course it is trivially true that the “posited reality” of a theory is as the theory describes it. Astrologers may believe that Brutus killed Caesar because he was a Sagittarius. This is a truth about the “posited reality” of astrological theory, a truth about what astrologers believe. If others—Marxist historians, say—believe that Brutus killed Caesar for some other reason, this is a truth about another, different “posited reality”.²²

Hence Quine seems here to assimilate truth to warranted belief. This indicates the real meaning of his suggestion, elsewhere, that “there is some underlying validity to the

---

²¹ Ibid.
²² This example is not entirely fair to Quine, since the theories of astrologers and astronomers can be plausibly seen as competing elements of a single, larger theory of the world. The alternative theories Quine has in mind, by contrast, differ more radically and globally. But this is not important at present.
correspondence theory of truth, as Tarski has taught us”. The underlying validity is that true sentences correspond to the “posited reality” of the theory to which they belong. Thus truth is purely infra-theoretical; there is no room for the question of whether the astrologer’s beliefs about the planets, or instead the astronomer’s, are really true. For that would be to suppose that we could we can make sense of the question of how our sentences or theories compare with reality as it is in itself. So it seems we must “settle for a relativistic theory of truth—rating the statements of each theory as true for that theory, and brooking no higher criticism”. And yet we can’t accept this conclusion, according to Quine, because “we own and use our beliefs of the moment, even in the midst of philosophizing”. When we say that a theory is true, we can only mean that corresponds to the “posited reality” of our own theory.

In some sense, then, the alternatives to our own theory are true; in another sense, they are all false—or at least, we have no choice but to believe that they are. Naturalism is supposed to resolve this paradox. The recognition (or seeming recognition) that the other theories are true in their different ways was merely epistemological. In reflecting on the slim basis of our own theory and the possibility of alternatives, we did not set aside “our beliefs of the moment”, our actual ontology and theory of the world. Epistemology is simply one chapter of that larger theory, and so cannot be intelligibly thought to cast doubt on it.

These are the only perspectives available to us: the perspective of the scientific study of scientific method, and the perspective of that same science. There is no third perspective from which we might consider how the epistemological claim that the other theories are true in their different ways bears on the ontological claims of our own theory. So the relativistic conclusion is cancelled by the fact that we cannot transcend our theory. Since we can’t, we take our “beliefs of the moment” to underwrite a kind of “higher criticism”. The suggestion appears to be that because we are so thoroughly confined to our own perspective, we have no choice but to believe that we are not.

Perhaps this conclusion is consistent, but it is unclear what it has to do with the question of whether our own theory alone, or all the alternative theories, or none, actually

---

23 Pursuit of Truth, 80.
24 Word and Object, 24.
25 Ibid., 25.
is true. At most, it might be objected, this rather tortuous reasoning suggests merely that we will inevitably believe that our theory alone is true. But then we might well wonder what it would mean to suppose that those alternative theories are true even though we don’t believe that they are. This view seems self-defeating, in the same way as an assertion of G.E. Moore’s paradoxical sentence “It’s raining but I believe it isn’t raining”. Even if it is consistent to say that (i) the alternatives to our own theory are all true and, also, (ii) we believe they are all false, it is impossible for us to consistently believe that assertion.

I will not attempt to evaluate Quine’s discussion of these confusing issues. What I want to stress is that, first, Quine’s remarks about “immanence”, “transcendence” and the like are not merely heuristic or inessential to his system. He clearly takes seriously the question of whether and how our thoughts and perceptions are shaped or limited by our conceptual scheme, and hopes to clarify it by appeal to science. Second, I want to stress that this not in any obvious sense a straightforward scientific question. Despite the superficial plausibility of Quine’s reply to Davidson, there is more to his project than a “logical analysis” or an empirical description of scientific methods and results. It is this additional, metaphysical element that lends philosophical interest to what would otherwise be a kind of highly speculative physiology or empirical psychology.

II

Quine’s influential criticisms of the analytic/synthetic distinction can be seen in this more philosophical light. The philosophical interest of analytic truth stems, in part, at least, from its connections with the Kantian issues considered in the previous section. Perhaps it is worth pointing out that there is some obvious and undeniable distinction between analytic and synthetic truth, and that Quine’s critique of analyticity is not intended to deny this. There obviously are such things as linguistic conventions, or at least there could well be. We might imagine, for example, that a committee is established with the function of determining proper English usage, and that it is decreed that “bachelor” and “unmarried man” are synonyms. In that case, it would be perfectly intelligible to say that it is true (simply) by convention that bachelors are unmarried men.
This is plainly very different from, say, the statistician’s discovery that bachelors are poor. No linguistic legislation of this kind could bring it about that bachelors are poor, although it might, of course, bring it about that things called “bachelors” are said to be “poor”. (For example, English speakers might be compelled to call fire hydrants “bachelors” and red things “poor”.)

The point is merely that there is a difference between saying that a is F and saying that “F” is true of the referent of “a”. The first claim is true just in case a is F; the second is true just in case certain linguistic conventions are in place, regardless of whether a is F. Even Quine concedes that there is an obvious and undeniable difference here, as he makes clear in later writings:

Analyticity undeniably has a place at a common-sense level, and this has made readers regard my reservations as unreasonable. My threadbare bachelor example is one of many undefeatable cases. It is intelligible and often useful in discussion to point out that some disagreement is purely a matter of words rather than of fact. The point can commonly be sustained and acted upon by paraphrase that circumvents a troublesome word.  

His contention is rather that there is no deeper, principled distinction to draw between conventional and factual truths, between knowing how to speak a language and knowing about extra-linguistic reality:

My reservations over analyticity are the same as ever, and they concern the tracing of any demarcation, even a vague and approximate one, across the domain of sentences in general … Elementary logic and the bachelor example are clear cases, but there is no going on from there.  

What Quine rejects, then, is the idea that there is a non-trivial and globally applicable concept of analyticity that might underwrite a more interesting distinction between the genuinely factual and the merely linguistic or conceptual. Logical empiricism was founded on the assumption that there are non-trivial analytic truths of this kind. The notion of analytic truth was to support a radical, deflationary solution of traditional philosophical problems—the problem of objectivity, of necessary truth, of

---

27 Ibid.
how a priori knowledge is possible. Thus it was thought that sentences such as “All possible objects of experience are spatio-temporal” might be shown to be similar in kind to “Bachelors are unmarried men”. If so, it would be simply confused to hold that our knowledge of such truths is knowledge of the necessary features of reality, or that it reflects the activity of the knowing subject. For it is not really knowledge at all, but merely an indication of the linguistic conventions in use.

Nor were the logical empiricists the only philosophers to deploy something like the analytic/synthetic distinction for philosophical purposes. P.M.S. Hacker invokes linguistic convention to explain the later Wittgenstein’s view that “logical possibility is relative to a language”:

What makes sense in a given grammar is dependent upon the conventions that constitute that grammar... Essences are reflections of forms of representation, marks of concepts, and thus made rather than found... In metaphysics we characteristically express our puzzlement concerning our forms of representation by factual questions about objects in the world instead of grammatical questions about language... We sense our conventions as limits, and so indeed they are—they are limits of sense. But that is not how they are taken by the metaphysician. For he projects our grammatical conventions on the world and then views them as de re necessities. This reading of Wittgenstein involves the most striking and puzzling ideas associated with the analytic/synthetic distinction. There are non-trivial, hidden linguistic conventions that represent “limits of sense”. Thus what seem to be mysterious features of reality are, in fact, merely features of “our form of representation”. (So, to borrow one of Wittgenstein’s examples, the sentence “Pain is a private sensation” is somehow akin to the sentence “Solitaire is a game I play by myself”.) There is some interesting sense in which possibility, necessity and truth are “relative to a language”—something more than the trivial observation that whether a sentence is true depends on what it means, which is of course partly a linguistic matter.

---

Notice that this line of thought is reminiscent of the view I ascribed to Quine in explaining his naturalistic defense of empiricism. The suggestion was that the methods and results of natural science are constitutive of what we take to be justification, rationality and reality; in some sense, the role of science is akin to that of logic. This is not surprising. Quine does not reject the idea that our linguistic conventions—or although he would not use the phrase—our “form of representation” shapes our view of the world. What he rejects is the idea that there is a sharp distinction between our conventions and the factual, empirical knowledge we represent according to those conventions. Thus, in a sense, he is a more radical conventionalist than Carnap or Wittgenstein. Logic, he thinks, is in some sense an empirical science; but by the same token, physics and geography are in some sense conventional. This is the source of the vacillation between the normative and the factual we noted earlier, in Quine’s response to radical skepticism.

There is a vast literature on this subject, and I intend only to suggest the implications of Quine’s disagreement with Carnap over the notion of linguistic convention. My point is merely that the interest of the distinction is closely connected to the metaphysical themes we have been considering. This is why Quine concludes from the failure of the analytic/synthetic distinction that

we cannot strip away the conceptual trappings sentence by sentence and leave a description of the objective world; but we can investigate the world, and man as a part of it, and thus find out what cues he could have of what goes on around him.

Subtracting his cues from his world view, we get man’s net contribution as the difference. This difference marks the extent of man’s conceptual sovereignty—the domain within which he can revise theory while saving the data.29

The analytic/synthetic distinction seemed to promise a way of isolating the purely conventional truths of our language from a purely factual “description of the objective world”. This method has turned out to be unworkable, but the question of conceptual sovereignty can be pursued naturalistically. We can “strip away the conceptual trappings” by clarifying, from within our current theory, the basic ontological

29 Word and Object, 5.
commitments of that theory. (As we noted earlier, Quine thinks that this will lead us to reject propositions, properties, mental states and much else.)

There is a sense, then, in which the analytic/synthetic distinction was a version of the scheme/content distinction, and seemed to promise a way of pursuing Kant’s question about the limits and objectivity of thought. Quine rejects that version of the distinction, but defends the underlying idea of a distinction between the objective world and the “conceptual trappings” by means of which we represent it. It is just this underlying idea that Davidson claims to find unintelligible: the idea that our conception of the world is determined partly by our conventions or concepts and partly by the world as it is apart from any conception of it. There is thus an important continuity between Davidson’s attack on the scheme/content distinction and Quine’s criticisms of the first two “dogmas” of empiricism. Quine’s suggestion was that the traditional distinction between the conceptual and the factual is untenable, and Davidson’s is that any such distinction is untenable. I’ll now explore some of Quine’s views in more detail, stressing what I take to be the central tension in his system.

III

Quine holds that all theories are radically under-determined by sensory evidence. In its more familiar form, the thesis of under-determination is simply that the evidence is insufficient to settle which of various hypotheses is true. Therefore, theory choice depends in large part on non-empirical considerations such as simplicity, or coherence with what seem to be independently warranted assumptions. This kind of under-determination does not suggest that alternative conceptual schemes are possible, or help to clarify what that would mean. On the contrary, if the evidence fails to determine which of two conflicting theories is true, both theories presumably share a significant range of concepts. Otherwise, there will be no way to identify a common subject matter concerning which they differ.

We can also put this point in linguistic terms: if certain sentences imply the negation of certain others, both sets of sentences must be expressible in a single language. Suppose that one theory implies that the earth orbits the sun, while another implies that
the sun orbits the earth. Each theory will at least make use of terms translatable by
“earth”, “sun”, “orbit”, “not” and presumably many others. (For surely it’s unlikely that
either theory implies just one sentence that can be translated into the vocabulary of the
other.) The difference between the two conflicting theories can be characterized, in this
case, by reference to a single, more comprehensive theory or language. If conceptual
schemes are, as Davidson and Quine agree, to be associated with languages, the
conclusion here must surely be that the two theories belong to a single conceptual
scheme.

Quine’s view, by contrast, is that the more comprehensive language or theory by
means of which this lesser kind of under-determination can be characterized is itself
radically under-determined by the evidence. In Word and Object, he writes that “the
truths that can be said even in common-sense terms about ordinary things are … far in
excess of any available data”:

This remains true even if we include all past, present and future irritations of all the
far-flung surfaces of mankind, and probably even if we throw in an in fact
unachieved ideal organon of scientific method besides… We have no reason to
suppose that man’s surface irritations even unto eternity admit of any one
systematization that is scientifically better or simpler than all possible others. It
seems likelier … that countless alternative theories would be tied for first place. 30
Even the notion of an enduring material object is theoretical. By the same reasoning,
Quine allows even that

some languages are perhaps so unlike ours that any translation of ‘there is’ or ‘ x’,
however cunningly contextual, would be far too far-fetched and Procrustean to rest
with. To entertain the notion of an ontology at all, known or unknown, for the
speakers of such a language would be an unwarranted projection on our part of a
parochial category appropriate only to our own linguistic circle. 31

Even “the true-false dichotomy itself” 32 is, in some sense, a “parochial” invention:

“Factuality, like gravitation or electric charge, is internal to our theory of nature”. 33

30 Word and Object, 22.
31 Pursuit of Truth, 28.
32 “Posits”, 244.
33 “Things”, 23.
The suggestion, then, is that the sensory evidence is so meager that even the most seemingly fundamental and indispensable concepts must be seen as theoretical constructs. Empirically equivalent theories may differ over basic ontology, over what objects and kinds of objects there are. Where there are such profound differences, it is hard to see how translation could be possible. How could typical English sentences, for example, be translated into a language free of any terms referring to enduring physical objects? Translation is at least in part a matter of correlating the terms and sentences of a language with a set of objects. Quine’s point is that there is no natural theory of objects by means of which translation between different languages is guaranteed. This line of thought appears to meet Davidson’s criteria for a difference in conceptual scheme. There is something neutral and common to radically different theories, the sensory evidence. Since this is not the common subject matter of the two theories, there is no paradox in supposing that they are not inter-translatable.

Perhaps we can further clarify the idea. Supposing that our own theory is true, we can conclude that our terms represent a certain “systematization” of sensory experience that corresponds to reality. Thus, we know that the predicate “electron”, for instance, is true of certain extra-linguistic (and extra-sensory) objects. Now suppose that there is another, radically different theory that imposes a radically different system on the same patterns of sensory experience. According to Quine, this theory may contain no terms that refer to the same objects as our own—no terms referring to electrons, sticks and stones, or anything else that we believe to exist. We begin to see why Quine holds that sentences are “about posited entities” and “are significant only in relation to a surrounding body of theory”. The alien predicates refer only to the imaginary objects of the alien theory, and never to anything real. And yet plainly they do refer, and so the alien sentences are meaningful in some basic sense. Its proponents will, for example, be able to offer a semantics for their language; they will be able to explain that their sentence “Fa” is true if and only if “F” is true of the referent of “a”.

Thus it seems we should distinguish between two varieties of reference, infra-theoretical reference and metaphysical reference. Infra-theoretical reference is a relation between a word and one or more of the “posited entities” that figure in the theory. It is in this sense that names of fictional objects or properties are said to refer, for example.
Metaphysical reference is a relation between a word and a real entity, posited or not. Translation from one language into another requires that the infra-theoretical reference relations of both languages can be correlated. But since the posited entities of certain theories are not real, there is no way to translate them into our own, or into each other. This would be like trying to correlate fictional infra-linguistic reference relations. For example, Nabokov writes in *Pale Fire* of an obscure central European kingdom, Zembla. It makes no sense to ask what place the term “Zembla” denotes in Tolkien’s Middle Earth, how far Zembla is from Oz, etc. Consequently, there are passages in *Pale Fire* that cannot be “translated” into the vocabulary of other fictions.

Of course in the case of fictional works there is some correlation, and it is typically quite extensive, and so it is often possible in principle to expand the vocabulary of one fiction so as to incorporate others. An ambitious writer might easily devise a story combining the fictional worlds of *Pale Fire* and *The Lord of the Rings*. This is because works of fiction that we can understand are recognizable manipulations of the “posited entities” of a single theory of the world. Even culturally or historically remote fictions still make reference to trees, oceans, cities, and the like. But perhaps this suggests the kind of semantic disparity Quine has in mind. It is as if any theory is a kind of elaborate fiction—“make-believe”, as Quine says—although luckily for us our own fiction happens to be a true description of the objective world. Since there is no guarantee that the infra-linguistic relations of empirically equivalent theories will be reducible to metaphysical reference relations, there is no guarantee that they can be correlated. Thus it would seem that Quine’s thesis of radical theoretical under-determination supports a kind of conceptual relativism: there are “countless” mutually incomprehensible ways of conceptualizing sensory experience. And perhaps, as we’ve seen, there is even a sense in which truth is relative to a language, theory or conceptual scheme. Either conclusion, however, seems to be flatly inconsistent with Quine’s theory of meaning. To see why, we’ll need to consider his theory of meaning in some detail.

Quine holds that meaning is by nature public, since in learning a language, we “depend strictly on overt behavior in observable circumstances”.34 (As we’ll see shortly, this is the root of the indeterminacy of translation.) In fact Quine’s theory of meaning is

---

34 *Pursuit of Truth*, 38.
considerable more complex. Just as sensory stimulation is the ultimate and only evidence for any theory, it is also the basis of linguistic meaning. Language acquisition is at bottom a neurophysiological process. The child or translator is conditioned to respond to patterns of sensory stimulation in socially acceptable ways. Dispositions to verbal behavior are enduring states of a speaker’s nervous system. Quine hopes to suggest, at least in outline, how the pre-theoretical notion of meaning can be eliminated in favor of empirical or stimulus meaning:

Empirical meaning is what remains when, given discourse together with its stimulatory conditions, we peel away the verbiage. It is what the sentences of one language and their firm translations in a completely alien language have in common.\(^\text{35}\)

While the details of Quine’s theory of meaning have changed considerably over time, the central idea that meaning ultimately consists in the “stimulatory conditions” of discourse has not. Ultimately, it would seem, this is simply a consequence of physicalism. Stimulatory conditions are physical conditions, describable in principle as “distributions of states and relations over elementary particles”.

Quine’s theory is roughly as follows. Observation sentences such as “That hurts!” or “That’s red!” are directly conditioned to stimulation. Competent speakers of English are disposed to assent to such sentences simply by the activation of certain nerve endings. The pattern of stimulation that causes or disposes a speaker to assent to an observation sentence is the “affirmative stimulus meaning” of the sentence. Other, more theoretical sentences, such as “That’s a fire truck” or “Fire trucks are red”, are conditioned to stimulation indirectly and holistically. Sensory stimulation alone is not enough to cause or dispose a speaker to assent to such sentences; but a group of theoretical sentences jointly imply observation sentences, and thus share in their empirical content. (Thus Quine holds that sentences of logic and pure mathematics, for example, are meaningful to the extent that they figure in empirical theories.) Insofar as we can speak of the meaning of an entire theory or language, then, it would seem that this

can only be the affirmative stimulus meaning of all the observation sentences held true by its proponents.

As Davidson and other critics have stressed, the concept of stimulus meaning seems to conflict with Quine’s fundamental thesis that linguistic meaning is a social art, that linguistic meaning is hence nothing over and above what a translator would be able to discern. Obviously it is impossible that two speakers undergo the same stimulation; if the meaning of ‘s’ is the pattern of stimulation that provokes utterances of ‘s’, it would seem that ‘s’ must have a different meaning for each speaker. And in any case, events occurring in a speaker’s nervous system are typically of no interest to a translator. He relies “strictly on overt behavior in observable circumstances”, whatever the neurophysiological facts may be. How then do private stimulus meanings issue in a learnable, public language?

At times Quine has speculated that there must be an extensive similarity of stimulus meanings across speakers. That is, we might suppose that the patterns of neural activation triggered in different people by red things, or hot things, are much the same; the only real difference is that each occurs in a different nervous system. But a fine-grained neural similarity across speakers is not particularly plausible, and more to the point, “such anatomical minutiae ought not to matter here”.\(^{36}\) Surely it is bizarre to suppose that the discovery of a great neural diversity across speakers of English or Japanese would reveal that they are not really communicating with each other.

Another solution rests on the assumption of a “pre-established harmony of standards of perceptual similarity, independent of intersubjective likeness of receptors or sensations”.\(^{37}\) Two people confronted with a fire truck and a red rose are innately disposed to rate their respective perceptions as similar, even though the sensory episodes thus associated by one speaker may be utterly unlike those associated by the other. (Of course, the shared standards of perceptual similarity may themselves be realized differently in different nervous systems.) Private stimulus meanings remain essential though, in that they are the objects of a speaker’s associative mechanisms. So, while private stimulus meanings are crucial to epistemology and to an account of language

\(^{36}\) Pursuit of Truth, 40.
acquisition, the socially conditioned structure that each speaker internalizes in his own way is all that really matters to linguistic meaning:

Different persons growing up in the same language are like different bushes trimmed and trained to take the shape of identical elephants. The anatomical details of twigs and branches will fulfill the elephantine form differently from bush to bush, but the outward results are alike.\(^{38}\)

The common structure, whatever its underlying mechanism in any particular speaker, is manifest in overt behavior and dispositions, and so accessible to the translator. It would seem that this leads to one of two conclusions. We might define empirical content only for individual speakers, characterizing a language simply as the public structure by means of which different speakers systematize their own patterns of stimulation. Or we might hold that this structural similarity, insofar as it results in a correlation of private stimulus meanings, determines the empirical content of a language. In that case the affirmative stimulus meaning of “That’s red” would be the set of patterns of stimulation in different speakers that prompt communal assent to the sentence. Likewise, the empirical content of an entire theory would be a set of patterns of stimulation in different speakers.

It isn’t clear which conclusion Quine draws, or indeed whether the view outlined above is quite what he intends. In any case, whatever the details, the empirical meaning or content of any theory will presumably be, if anything, a certain set of neural events\(^{39}\). This is why Quine at times defends a kind of verificationism. Since the “meaning” of a theory is simply its empirical content, it is incoherent to suppose that a theory might predict every possible observation and nevertheless be false:

Our scientific theory can indeed go wrong, and precisely in the familiar way: through failure of predicted observation. But what if, happily and unbeknownst, we have achieved a theory that is conformable to every possible observation, past and future? In what sense could the world then be said to deviate from what the theory claims? Clearly in none… Our overall scientific theory demands of the world only that it be so structured as to assure the sequences of stimulation that our

\(^{38}\) *Word and Object*, 8.

\(^{39}\) Maybe this must be a set of both actual and possible neural events; but I can’t see that this complication, in light of Quine’s aversion to quantifying over possibilities, is important in the present context.
theory gives us to expect. More concrete demands are empty, what with the freedom of proxy functions.\textsuperscript{40}

The world cannot “deviate from what the theory claims” because the theory really makes no claims about the world at all, except perhaps the vacuous claim that the world is structured so as to produce in us certain “sequences of stimulation”. Indeed, strictly speaking, the theory claims \textit{nothing}—not even that those sequences occur. It is simply a mechanism by means of which we are caused to anticipate future sequences of stimulation on the basis of past sequences.

This is hard to reconcile with Quine’s contention that theories are radically under-determined by the sensory evidence. Suppose we have two empirically equivalent and empirically \textit{ideal} theories: both are “conformable to every possible observation, past and future”. The present conclusion is that the world cannot “deviate” from what an empirically ideal theory says about it, for the theory has no content over and above its empirical content. But, by the same token, neither can either of these theories deviate from what the other says. Given Quine’s empiricist theory of meaning, any two theories conditioned with equal strength to the same set of neural events are \textit{synonymous} (“stimulus synonymous”). What is it that is supposed to be under-determined by the sensory evidence, then? Unless the two theories differ in meaning, there is nothing that the evidence fails to determine, no question about the world that remains unanswered. If they do differ in meaning, it must be a difference that transcends \textit{empirical} meaning.

Quine is unsure how to resolve this problem, and sometimes falls into blatant contradiction. Consider one of his last discussions of empirical equivalence, in \textit{Pursuit of Truth}. We are to imagine two empirically equivalent theories. One rests on “our commonsense conception of infinite space and rigid bodies that move freely without shrinking or stretching”, the other on “the conception of a finite spherical space in which

\textsuperscript{40} “Things”, 22. A “proxy function” is a one-to-one transformation that systematically shifts the referents of an entire language without disturbing the truth values of its sentences. Quine takes the existence of such functions to be a proof of referential indeterminacy: there is no “fact of the matter” as to what the objects of any theory are. For example, we can stipulate that English terms will now refer to their “cosmic complements”: each term will now refer to everything in the universe except its customary referent. This will not change the truth-values of any sentence. Quine takes this to show that reference, and hence ontology, are “indeterminate”: the claim that ‘a’ refers to a or that ‘F’ is true of a has no truth-value, except relative to an arbitrary assignment of reference relations.
those bodies shrink uniformly as they move away from [the] center”. 41 There is no way to convert one theory into the other, as the term “center of space” figures essentially in the one and has no counterpart in the other. Quine writes that the two theories embody different but equally defensible “ways of conceiving the world”. 42 For they plainly differ “more deeply than in mere choice of words”. And yet, ultimately, “the cosmic question whether to call two such systems true … simmer[s] down, bathetically, to a question of words”. 43 This is merely a matter of linguistic “convention”, as “the rival theories describe one and the same world”. 44

IV

Quine’s thesis of the indeterminacy of translation offers another argument for the possibility of alternative conceptual schemes. The thesis is that any natural language can be rendered equally well by mutually incompatible manuals of translation; these represent different ontologies or conceptual schemes. Thus Quine writes that in pairing the sentences of a newly discovered language with his own, the translator must rely on the “conjecture” that

the native conceptual scheme is, like ours, one that breaks reality down somehow into a multiplicity of identifiable and discriminable physical things, be they rabbits or stages or parts.

And yet this may be a mistake, since

the native attitude might, after all, be very unlike ours… The native point of view might be so alien that from it there would be no semblance of sense in speaking of objects at all, not even of abstract ones like rabbithood. Native channels might be wholly unlike Western talk of this and that, same and different, one and two. 45

In these passages, the indeterminacy of translation is difficult to distinguish from the thesis of radical theoretical under-determination. The native conceptual scheme might be “wholly unlike” our own in that ours “breaks reality down” in a certain way

41 Pursuit of Truth, 96.
42 Ibid., 102.
43 Ibid., 101.
44 Ibid.
45 “Meaning”, 154.
while the other may break it down differently—or perhaps not at all, Quine suggests. Our own language is one “systematization” of sensory stimulation among others. Again, the claim is that the evidence is too slight to guarantee a common ontology by means of which translation can proceed. In translating the native language into English, we project our own ontology and scheme on to it; but their language might, undetectably, represent an utterly “alien” view of the world. The “indeterminacy” of translation, then, seems to be a misnomer. Translation and linguistic meaning are not indeterminate, but simply under-determined. Given the constraints of translation, there is no way to find out whether the native scheme is like our own or entirely alien. This would seem to be a direct consequence of the under-determination thesis, and not, as Quine typically claims, a different and additional thesis.

Let us call this the epistemological reading of the indeterminacy of translation. This is not, though, Quine’s official account of indeterminacy. It would be obviously invalid to conclude that two translations capture the scheme or ontology of the natives equally well from the premise that there is no way to know which one is right. Rather, Quine officially holds that translation is ontologically—that is, genuinely—indeterminate. This is a consequence, or perhaps a presupposition, of the theory of meaning considered in the previous section. The argument for indeterminacy begins from the assumption that “what is utterly factual” in translation, “is just the fluency of conversation and the effectiveness of negotiation that one or another manual of translation serves to induce”. 46

Since dispositions to overt behavior exhaust the cognitive meaning of a language, there is an enormous freedom for invention and conjecture in assigning truth conditions to the sentences being translated. The gap between the facts of meaning and the interpretations the translator places on them is so wide that two equally good translations of the same sentence may not be interchangeable in the target language. Indeed, the target language sentences paired by rival manuals with a single source language sentence may differ in truth-value. The same is true of our interpretation of our own words, since we may in principle translate English into a foreign language according to one manual of translation and back into English according to another. (Notice that if this is right, the

46 Pursuit of Truth, 43.
scenario of the empirically equivalent theories of the world is even more mystifying than it seemed. Not only is it unclear whether the two theories are genuinely different, representing different conceptions of the world. If they do differ in this way, it remains indeterminate how they differ, as each theory is expressible by mutually incompatible translation manuals. And if they don’t differ, but are merely different formulations of a single theory, it remains indeterminate which single theory they both express.

It is hard to imagine the alleged indeterminacy in any detail, and Quine cannot offer any examples. However, the more limited phenomenon he calls “the inscrutability of reference” is supposed to suggest how it might arise. (Unlike the indeterminacy of translation, Quine thinks that this can be proven by the proxy functions mentioned earlier.\(^{47}\)) Quine’s original example concerns the question of how to translate an imaginary language, Jungle, that has never been translated into English or any other language that has. Suppose that the Jungle word “Gavagai” is uttered when English speakers would say “rabbit” or “There’s a rabbit”. The translator will conclude “that a gavagai is a whole enduring rabbit”, but “he is just taking for granted that the native is enough like us to have a brief general term for rabbits and no brief general term for [temporal] rabbit stages or parts”.\(^{48}\) For all he knows, “gavagai” might refer only to temporal stages of rabbits or to something else entirely.

Again, the inscrutability of reference sometimes seems to be, as the name suggests, an epistemological thesis. But Quine’s official position is that there is nothing to know about the reference of alien terms that is not open to the translator. If we construe “Gavagai” as a sentence, perhaps this can be seen as a case where two translations of the same sentence have different truth-values. A temporal stage of a rabbit is not the whole rabbit, and so it can’t be true that some particular thing observed by the native is both a rabbit and a rabbit stage. But, supposedly, there may be a translation of “Gavagai” that assigns it the first meaning and another, equally good translation that

---

\(^{47}\) For example, we can stipulate that English terms will now refer to their “cosmic complements”: each term will now refer to everything in the universe except its customary referent. This will not change the truth-values of any sentence. Thus the sentence “Rabbits are furry” will now be interpreted to mean that complements of rabbits are complements of furry things; this is true if and only if rabbits are furry. Quine takes this to show that reference, and hence ontology, are “indeterminate”: there is no fact of the matter as to whether ‘rabbits’ is true of rabbits rather than everything in the universe other than rabbits, except relative to an arbitrary choice of reference scheme.

\(^{48}\) “Meaning”, 153.
assigns it the second. And again, the same is true of the English sentence “There’s a rabbit”. The indeterminacy of translation, therefore, is an indeterminacy of meaning generally, infecting the translator’s language as much as the language being translated. Hence, as Quine puts it, there is no fact of the matter as to what we are talking about or what we are saying.

As Dorit Bar-On has argued, this astonishing conclusion seems to invoke the very conception of meaning that it rules out. We are to imagine that the translator is applying a certain set of “analytical hypotheses” to the source language, rather than some other. This is to presuppose that the rival manuals of translation are different. For example, we are to imagine that the translator’s sentence “A gavagai is a rabbit” means that a gavagai is a rabbit, and not—according to a second set of analytical hypotheses—that a gavagai is a rabbit stage. What accounts for this difference in meaning between analytical hypotheses? There is no difference as regards the linguistic meaning of the source language, i.e., the dispositions and overt behavior of its speakers. This is why translation is supposed to be indeterminate.

But for the same reasons, it would appear that there is no difference here as regards the linguistic meaning of the target language, since it is equally indeterminate whether the English term “rabbit”, as used by the translator, refers to rabbits, rabbit stages, recurring instances of a universal, or something else. But then is there any fact of the matter as to whether the meaning of the Jungle sentence is indeterminate? If not, the indeterminacy thesis is false. As with the thesis of radical under-determination, then, we may wonder what is not determined. The problem is that Quine’s conclusions are of no interest except if they are understood on the basis of the various intuitive semantic notions that they are supposed to show to be baseless. This is what I will try to show in the next section. The upshot, I think, is that we do not have any clear understanding of Quine’s claims about semantic indeterminacy or conceptual relativism.

---

49 Dorit Bar-On, “Indeterminacy of Translation – Theory and Practice”, *Philosophy and Phenomenological Research*, 53 (1993): 804. “[Quine] has tried to construe what we take to be linguistic facts in terms of … what we take to be linguistic facts! For a proponent of indeterminacy, the facts that the relativist construal [of semantics] adduces in its attempt to explain why we take things to be as we do (in ordinary translation) should be just as problematic as the facts to be explained away by the construal. They belong to the very same class of facts whose existence the indeterminacy thesis denies.”
Let us consider more closely Quine’s position regarding the supposed indeterminacy of translation. Quine’s suggestion is that the difference between the manuals will be discernable to English speakers in that the first pairs the Jungle sentence \( r \) with an English sentence \( s \), and the second with an English sentence \( t \), where \( s \) and \( t \) are not interchangeable. Indeed, English speakers might count one of these sentences as true, and the other false, and yet also come to see that either is an equally good translation of the Jungle sentence. Presumably \( s \) and \( t \) must then differ in meaning, at least in the sense that their truth-conditions are different. For otherwise we have no way of knowing what it means to say that they might be assigned different truth values. We should assume, then, that \( s \) is true if and only if \( s \), and that \( t \) is true if and only if \( t \).

But are those in fact the truth-conditions of \( s \) and \( t \)? If we can with equal warrant assign the Jungle sentence \( r \) the truth-conditions of either \( s \) or \( t \), we can also translate \( s \) into Jungle according to one manual and back into English according to the other. Suppose that the first manual of translation for Jungle, \( J_1 \), yields as a translation of the Jungle sentence \( r \)

(i) \( r \) is true in Jungle if and only if \( s \),

while the second, \( J_2 \), yields instead

(ii) \( r \) is true in Jungle if and only if \( t \).

If so, we can translate the English sentence \( s \) that appears on the right-hand side of (i) into Jungle by

(iii) \( s \) is true in English if and only if \( r \),

and translate \( r \) back into English by (ii). But if \( s \) is true in English if only if \( r \), \( r \) is true in Jungle if and only if \( s \), and \( r \) is true in Jungle if only if \( t \), then \( s \) is true in English if and only if \( t \). So according to one manual of translation for English, \( E_1 \),

(iv) \( s \) is true in English if and only if \( s \).

But according to another, \( E_2 \),

(v) \( s \) is true in English if and only if \( t \).
Thus, if it is indeterminate whether the meaning of Jungle sentences is best captured by J1 or instead by J2, it is equally indeterminate whether the meaning of English sentences is best captured by E1 or instead by E2. In particular, it is indeterminate whether the meaning of $s$, as used by the translator, is best captured by (iv) or instead by (v). Hence in the absence of an arbitrary choice of manuals for English, it is indeterminate whether J1 assigns $r$ the same truth-condition as J2. But it was precisely this kind of semantic relativity that, according to Quine, reveals the meaning of Jungle sentences to be indeterminate. By the same reasoning, then, it is also indeterminate whether J1 and J2 are different manuals for Jungle, or merely notational variants of a single manual. But if it is indeterminate whether they are different manuals, then it is, in particular, not true that they are. Then it is indeterminate whether the linguistic meaning of Jungle is indeterminate, and so it is not true that the linguistic meaning of Jungle is indeterminate.

This objection to Quine’s reasoning might be expressed as a dilemma. Quine’s conclusion might be either of the following:

(i) Translation is indeterminate in relation to sentences, in that different sentences in one language can be equally good translations of a single sentence in some other.

(ii) Translation is indeterminate in relation to the meanings or truth conditions of the sentences under translation: the same sentence can be assigned different meanings or truth conditions by empirically equivalent manuals of translation.

The first claim is consistent with Quine’s rejection of any facts of meaning that transcend the behavioral and physiological facts. But this would trivialize the indeterminacy thesis, since it does not leave room for any real difference between manuals of translation, except differences to do with the sounds and shapes of the marks and noises that “translate” one language into another.

The second claim says something extremely radical and interesting, but violates Quine’s austere conception of linguistic meaning. Moreover, it renders Quine’s argument self-defeating. For if we say that there is some fact as to, say, the truth conditions assigned to a sentence by rival manuals of translation, this requires that the sentences in the meta-languages used to formulate those manuals have determinate truth conditions. But if those languages contain sentences with determinate truth conditions,
there is no reason to deny that the language under translation does as well. So, on this horn of the dilemma, there is presumably a fact of the matter as to which manual (if any) correctly translates the alien language, even if there is no way of discovering this kind of fact empirically.

In Quine’s defense, we might emphasize that while E1 can be converted into E2, this does not show that \( s \) and \( t \) are inter-changeable in any particular version of English. (Nor does it show that they are inter-changeable in some more general sense. Considered merely as strings, abstracted from all particular assignments of truth conditions, \( s \) and \( t \) can’t be intelligibly said to be inter-changeable, or not.) The suggestion, then, is that the indeterminacy thesis does not presuppose that there is any particular semantic difference in relation to which the meaning of the Jungle sentence \( r \) is indeterminate. It is the thesis that the meaning of \( r \) is indeterminate in relation to some semantic difference or other. For any version of English, there is some semantic difference or other between \( s \) and \( t \), and so there is some such difference in relation to which the meaning of \( r \) is indeterminate.

This may be true, but if so, as suggested just now, we should assume that the situation is the same as regards Jungle. If the native language in use is Jungle as (homophonically) translated by J1, it is perfectly determinate that the truth-condition of \( r \) is that \( s \), and not that \( t \); if it is Jungle as translated by J2, the truth-condition of \( r \) is that \( t \), and not that \( s \). So if this line of argument shows that there is a difference between J1 and J2, due to the difference between E1 and E2, it also shows that there is at most one accurate manual of translation for Jungle.

VI

Quine suggests that the semantic relativity can be resolved because we can choose as our manual of transformation the identity transformation, thus taking our home language at face value… Reference is then explicited in disquotational paradigms analogous to Tarski’s truth paradigm: ‘rabbit’ thus denotes rabbits, whatever they are, and ‘Boston’ designates Boston.\(^{50}\)

\(^{50}\) *Pursuit of Truth*, 52.
But it is hard to see how this resolves anything. Certainly the quoted and disquoted expressions are syntactically identical, but this in itself fails to establish that the first designates Boston rather than something else. Whether this “disquotational paradigm” is in fact true just in case ‘Boston’ designates Boston depends on which version of English is in use as the meta-language; but this is just what Quine contends is indeterminate. What can we make of this?

Christopher Hookway defends Quine’s appeal to the “face value” semantics of our language by distinguishing ordinary, unreflective talk from philosophy:

Such semantic notions—rules and translation manuals—have no role in ordinary speech. In deciding whether to assent to a sentence, I respond ‘blindly’ with no sense of rules which force me to respond in one way rather than another. Philosophical reflection does not recover semantic facts which are unconsciously at work in normal speech behavior… So, we can think of ordinary speech about the physical world as pre-translational... Ordinary talk of truth and reference is not guided by sophisticated theory but rests simply upon disquotational principles. The embeddedness of such principles grounds everyday claims about what must exist for our theories to be true. This is why it is supposed that our ship can be securely afloat while we hold determinate translation relations at bay.51

Hookway’s view, then, is that, for Quine, homophonic translation is mere disquotation. It does not involve the assignment of a particular scheme or ontology, while more perverse assignments of truth conditions involve us in a more “sophisticated theory”. But what exactly is the difference? Certainly it seems reasonable to suppose that

(R) “Rabbit” is true of \( x \) if and only if \( x \) is a rabbit

is not, as it stands, a translation of the word “rabbit”. (Setting aside the most obvious reason: a “translation” of an English word by that same English word would be of no use to someone who doesn’t already speak the language.) Intuitively, a translation of a term or sentence conveys its meaning. And I can understand (R) without knowing what the word “rabbit” means, just as I can understand

(S) “Schmabbit” is true of \( x \) if and only if \( x \) is a schmabbit.

Indeed, (S) is not only intelligible, but true—if there is such a thing as a schmabbit, the word “schmabbit” is true of it, and if “schmabbit” is true of anything, it is true of schmabbits. Plainly it is not a translation of the word, though. Not only does a person’s knowledge that (S) is true fail to give him any knowledge of the meaning of “schmabbit”, the word has no meaning and thus can’t be translated at all.

Suppose then that our ordinary speech is pre-translational in this sense. In what way could disquotation, so understood, determine that “Rabbits are mammals” means that rabbits are mammals, and not that rabbit stages are stages of mammals? By hypothesis, disquotation is neutral with respect to either proposition, and does not even guarantee that “Rabbits are mammals” is any more meaningful than “Schmabbits are schmammals”. If the “face value” construal of our own language amounts to no more than this, it has no bearing on the question of what, if anything, our words mean. But if it does bear on that question, surely homophonic translation must be as “sophisticated” as its more perverse counterparts.

Even assuming, though, that our ordinary use of words somehow settles the semantics of our own language, and that this can be explained without violating Quine’s theory of meaning, the indeterminacy thesis is still self-defeating. For presumably whatever it is that we do in “taking our home language at face value” can be done by speakers of other languages as well. (Or at least, some further argument not provided by Quine or Hookway is needed here to explain why we alone are able to do whatever this is.) They need only take their language “at face value” to resolve the semantic relativity of their language. This would determine which of the various incompatible manuals of translation for Jungle, if any, is the right one.

So, to restate the dilemma noted earlier, the “face value” construal of our own language either does or does not involve the assignment of determinate truth conditions (to our own sentences). If it does not, there is nothing in relation to which our translations of other languages are indeterminate, since the different sentences that we can use to translate a single sentence in some other language are themselves indeterminate, and fail to conflict. On the other hand, if the “face value” construal does involve an assignment of truth conditions, with the result that we can take certain of our
own sentences to differ in truth value, there is no reason to maintain that other languages
are indeterminate.

Presumably for these reasons, Quine is deeply ambivalent as regards the idea of a
conceptual scheme. Even where he dismisses the idea, he also seems to appeal to it:
Can we even imagine any basic alternative to our own object-positing pattern?
Perhaps not; for we would have to imagine it in translation, and translation imposes
our pattern. Perhaps the very notion of such radical contrast of cultures is
meaningless, except in this purely privative sense: persistent failure to find smooth
and convincing native analogues of our own familiar accessories of objective
reference, such as articles, the identity predicate, the plural ending. Only by such
failure can we be said to perceive that the native language represents matters in
ways not open to our own.\footnote{52 “Meaning”, 155.}

The “very idea of such radical contrast of cultures” comes to little more than the idea that
some languages may be very awkward to translate. But then what is there to “impose” in
translation, and on what? This is the same vacillation we noted earlier, in connection
with the scenario of empirically equivalent theories of the world. As before, Quine says
that the difference is merely verbal, a matter of linguistic convention; but he also says
that it is a matter of incomparably different schemes, ontologies, ways of conceiving the
world. There is something to discover, but we will never be able to discover it.

Often these two inconsistent claims appear side by side. In finding that the native
is “much like us”, we may only have done a “thorough job of reading our own provincial
modes into the native’s speech”:
Usener, Cassirer, Sapir and latterly B.L. Whorf have stressed that deep differences
of language carry with them ultimate differences in the way one thinks, or looks
upon the world. I should prefer not to put the matter in such a way as to suggest
that certain philosophical propositions are affirmed in the one culture and denied in
the other. What is really involved is difficulty or indeterminacy of correlation. It is
just that there is less basis of comparison—less sense in saying what is good
translation and what is bad—the farther we get away from sentences with visibly
direct conditioning to nonverbal stimuli and the farther we get away from home
ground.\textsuperscript{53}

It seems, accordingly, that Quine’s theory of meaning is inconsistent with both the under-
determination of theory and the indeterminacy of translation. His efforts to explain what
it is that is under-determined by sensory evidence or indeterminate in relation to the facts
of meaning seem to be little more than gestures towards notions that have no place in his
austere system—“provincial modes”, “native channels”, “ontologies”, “world views”,
“conceptual schemes”. As we noted earlier, Quine suggests that Davidson reads too
much into his occasional use of such phrases; they play no essential role in his
philosophy. We can see now that this is not true. They have no legitimate role, perhaps,
but they are essential to his system, and to his larger project of gauging the extent of our
“conceptual sovereignty”.

\textbf{VII}

Quine does not anywhere address this problem directly. He seems not to
distinguish between the problem of semantical skepticism or nihilism and the problem of
epistemological skepticism. (Perhaps this is another example of his refusal to distinguish
between semantics and epistemology.) Consider, for instance, his admission that his
theory of knowledge “arouses certain logical misgivings”:

Is not our very talk of light rays, molecules, and men then only sound and fury,
induced by irritation of our surfaces and signifying nothing? The world view
which lent plausibility to this account of our knowledge is, according to this very
account of our knowledge, a groundless fabrication.\textsuperscript{54}

The original question was whether our talk of external things is nothing more than “sound
and fury, signifying nothing”, not whether our theory of the world is a “groundless
fabrication”. Groundless fabrications are, of course, meaningful; they mean something
more than what the evidence supports. The same is true of any theory, if Quine’s radical
formulation of under-determination is true. The question is how this is intelligible if

\textsuperscript{53} Ibid., 172.
\textsuperscript{54} “Posits”, 229.
linguistic meaning is what Quine claims: How can Quine make literal sense of the claim that we have a “world view”, “theory of the world”, “ontology” or “conceptual scheme”?

A superficial answer to this question might take essentially the same form as Quine naturalistic response to epistemological skepticism. For as we saw earlier, there is a similar tension in his system, between his seemingly skeptical epistemology and his unqualified realism about our current theory of the world—between empiricism and naturalism. If our theory is as poorly supported by sensory evidence as Quine claims, why should believe that it is true? Quine’s response is in part that skepticism is an empirical hypothesis; for example, it is an empirical hypothesis that China does not exist, or that there are no elm trees. It is based in empirical evidence of optical illusions, hallucinations and the like; so it is fair to appeal to science in response. Science reassures us that the world is largely what it seems to be.

But Quine sometimes makes use of a different argument, also suggested earlier. Our current methods and theory are constitutive of what we take to be knowledge, truth, justification and reality. It makes no sense to ask whether our theory is really true, or warranted, because it is only “within” that theory that those notions have any intelligible application:

We cannot significantly question the reality of the external world, or deny that there is evidence of external objects in the testimony of our senses; for, to do so is simply to dissociate the terms ‘reality’ and ‘evidence’ from the very applications which originally did most to invest those terms with whatever intelligibility they may have for us.\(^\text{55}\)

Now perhaps we might argue, similarly, that if we have reason to accept Quine’s account of meaning, we have equal or greater reason to believe the larger theory of the world of which it is a chapter. And that theory comprises not only such sentences as, say, “There are stars” or “Stars are hotter than ice”, but also such sentences as “The word ‘star’ is true of stars” and “The sentence ‘Stars are hotter than ice’ is true if and only if stars are hotter than ice”. These are perhaps less basic to our overall theory of the world than the sentence “There are stars”, but they are surely far more basic and obviously true than the

\(^{55}\) “Posits”, 229.
empiricist theory of meaning. So we can only conclude that our theory is nothing but a device for anticipating sensory stimulation and that, somehow, it is also a conception of the objective world, true or false depending on how things are in the objective world.

However this is not a very satisfying response in the absence of any account of how these seemingly contradictory conclusions could both be true. Maybe the larger conclusion is that our beliefs are bound to be hopelessly incoherent. This pessimistic conclusion was less compelling as regards the tension between Quine’s seemingly skeptical epistemology and his realist ontology. Perhaps it seems very improbable, intuitively, that our beliefs could be largely true given the slight evidence in their favor; but it seems flatly incoherent to say that a pattern of physiological responses and dispositions could constitute a set of beliefs about the objective world. This doesn’t seem to be the kind of thing that could be meaningful, warranted or unwarranted, true or false.

Here I agree with Hans-Johann Glock: Quine’s theory of meaning is not a kind of relativism, but “semantic nihilism: if there is no fact of the matter as to what we mean, then there is no such thing as linguistic meaning or understanding, and hence no meaningful communication, discourse or argument”. Indeed, as we’ve seen, it would seem to be equally impossible to explain even the appearance of linguistic meaning, communication, understanding, discourse or argument except by appeal to the idea of propositional meaning. So although Quine seems to deny that there is any such species of meaning, most of what he says is intelligible only on the assumption that there is—whatever exactly its nature may be.

The idea that language represents a “systematization” of sensory experience can no doubt be naturalized, as Quine would have it. Language acquisition and theory construction plainly are, at some level, behavioral and ultimately neurophysiological processes. The question is how this fact can be intelligibly taken to imply any conclusions about “man’s conceptual sovereignty” or anything of the sort. We have not yet seen how a description cast in these radically impoverished terms could serve even to explicate the concept of a belief or a theory, let alone the more puzzling claim that “we

---

must speak from within a theory”, or that there could be incommensurably different beliefs or theories.

It seems unlikely, then, that Quine’s talk of conceptual schemes can be explained away as merely heuristic or figurative. Davidson is right to say that Quine is committed to some form of conceptual relativism. At the very least, he appears to be committed to the possibility of theories or languages so alien that we would be unable, even in principle, to understand them. This would appear to follow from Quine’s radical version of the under-determination thesis, and it also appears to be implied by the thesis of the indeterminacy of translation. We’ve seen that this line of thought is in tension with other positions that Quine defends. In particular, it appears to be inconsistent with his radical skepticism concerning the very notion of meaning. Perhaps it would be more accurate, then, to say that some of Quine’s basic positions, which may or may not be consistent with the rest of his philosophy, provide a basis for a distinction between conceptual schemes and empirical content.

VIII

It is interesting to note that Davidson has argued against Quine’s doctrine of the indeterminacy of reference in essentially the same way as I have argued against the indeterminacy of translation. The argument might be taken to show the emptiness or unintelligibility of conceptual relativism. In effect, Davidson thinks that Quine’s semantic skepticism or nihilism does not go far enough. Quine rejects intuitive semantic notions, but he seems to think, at times, at least, that there can be perspectival semantic facts, so to speak: semantic facts relative to a manual of translation, for example.

Davidson correctly points out that this position is self-defeating: Suppose we could fix the ontology of ‘refers’ by relativizing it. Then we would have fixed the ontology of the language or speaker we were using the word ‘refers’ to characterize. It may be said: but the fixing is only relative to an arbitrary choice [between empirically equivalent manuals of translation]. That choice is not dictated by any relevant evidence. Hence the inscrutability [of reference]. This reply misses the point of the difficulty. The fixing of reference and ontology for
the object language has been done on the basis of an arbitrary choice; but the arbitrary choice succeeds in doing this only if the relativized ‘refers’ of the metalanguage has been nailed down.\textsuperscript{57}

This objection to the indeterminacy of reference has the same form as my objection, explained earlier, to the indeterminacy of translation. Unless there are determinate semantic facts about the language we are using to discuss the semantics of some other language, there is no sense in which the semantics of the second language is indeterminate, i.e., determinate only relative to some particular theory, translation or interpretation of the language under discussion.

If we say that “Quine” refers to Quine, relative to a certain account of English, this claim about (relative) reference is true only if, for example, the term “Quine”, as used in making this claim, \textit{in fact} refers to some particular thing and nothing else. If that occurrence of “Quine” does not determinately refer to a unique object, no determinate claim about reference relative to some theory or manual of translation has been made. And Quine’s official view, at least, is that no occurrence or use of any word is ever determinate in this sense, except relative to one among various equally adequate manuals of translation. Davidson concludes that, if we take seriously Quine’s doctrine of the inscrutability or relativity of reference, “any claim about reference, no matter how many times relativized, will be as meaningless as ‘Socrates is taller than’”.\textsuperscript{58}

Davidson compares the problem in coherently formulating the doctrine of referential or ontological relativity to the problem of formulating claims of cultural relativism. He says little about the similarity, but presumably the problem is just that in saying that truth is relative to a culture, for example, one would seem to be in possession of a concept of truth that is \textit{not} relative to a culture—i.e., whichever concept it is that enables us make claims of the form “s is true relative to culture C”, and take them to be true. If such claims are ever true, it is false that truth is relative to a culture. For suppose that my claim that some sentence \(p\) is true relative to (say) Chinese culture is itself true only relative to my own (say) North American culture. In that case, I have not succeeded even in saying that something is true relative to Chinese culture, but only that, instead, it

\textsuperscript{57} Donald Davidson, “The Inscrutability of Reference”, in \textit{Inquiries into Truth and Interpretation}, 232.

\textsuperscript{58} “Inscrutability”, 234.
is true relative to some other culture that $p$ is true relative to Chinese culture. But *that* could be true even if $p$ is not true relative to Chinese culture.

This is the point of Quine’s remark, which Davidson cites approvingly: one cannot “proclaim cultural relativism without rising above it”, but one cannot “rise above it without giving it up”.\(^{59}\) The parallel as concerns the supposed relativity of reference is that one cannot say what it is that is indeterminate in reference, except relative to a manual of translation, except by making use of referring terms that uniquely pick out certain entities. But there is more than a parallel here, of course, since, as we’ve seen, Quine often argues from semantic indeterminacies to a kind of conceptual relativism. Or, more precisely, he argues from the impossibility of ever discovering a unique set of reference or truth conditions for an alien language to the conclusion that the possibility that the alien language “represents matters in ways not open to our own”,\(^{60}\) that it is associated with an alternative conceptual scheme.

So Quine appears to hold, albeit ambivalently, that his radical position regarding the nature of linguistic meaning and communication allows for the possibility of differences in conceptual scheme. I’ve argued that this is not a consistent position. Davidson’s objection to Quine’s thesis of the indeterminacy of reference amounts to the same kind of objection: to be consistent, Quine should take the even more radical position that all talk of reference is meaningless. (The same would seem to be true regarding any talk of any kind of putative semantic facts, of course.) On this more consistent and still more radical position, the very idea of a difference in conceptual scheme would be strictly unintelligible. For on this view, there just are no semantic facts of any kind, no facts about how language is related to the world. And if so, there cannot, obviously, be facts about the *different* ways in which languages are related to the world, i.e., facts about different forms of representation, different ways of organizing experience, and so on. But it is hard to make sense of the idea that there are no such facts—at least as hard, I would argue, as it is to make sense of conceptual relativism.

Let us briefly summarize the foregoing discussion. My chief aim in this chapter has been to illustrate, by reference to a fairly detailed discussion of Quine’s philosophy,


\(^{60}\) “Meaning”, 155.
how one might develop and defend a theory of conceptual schemes. In exploring Quine’s
treatment of the scheme idea, we’ve also noted some important tensions internal to his
characterization of conceptual schemes that may be symptomatic of more general
tensions in the very idea of a conceptual scheme.

I’ve also argued that Quine’s discussions of semantic indeterminacy and, by
extension, the possibility of a difference in scheme cannot be given a satisfactory
interpretation. Either his views ultimately appeal to the semantic notions that he claims
to reject, or else he consistently rejects those notions, depriving himself of the conceptual
resources needed to state his radical claims. Davidson seems to object to at least some of
Quine’s doctrines on similar grounds; it may be that his rejection of the scheme idea is, in
part, an effort to pursue Quine’s basic philosophical commitments further than Quine
himself. It may be that his interest in the scheme idea stems partly from reflection on
these difficulties in interpreting Quine’s position regarding ontological and, by extension,
cultural relativism. But I have not tried to argue for this last suggestion. In light of these
considerations, I’ll now turn to the task of interpreting and evaluating Davidson’s
argument against the intelligibility of the scheme idea.
CHAPTER 2
A VERIFICATIONIST ARGUMENT?

Many commentators have agreed with Richard Rorty’s early assessment that Davidson’s discussion of conceptual schemes “is verificationist, and turns on the recognizability of persons using a conceptual framework different from our own”. I do not think that the argument, so construed, is at all plausible. Since this is a common reading, though, it is worth seeing why such an argument fails to establish that the very idea of a conceptual scheme is unintelligible. In later chapters, I will set out and evaluate a different reading of Davidson’s argument that I find more powerful and interesting.

I

Before turning to Rorty’s discussion, which will be my focus here, it is perhaps worth briefly considering a reading of the argument that attributes to Davidson an even more extreme form of verificationism. Simon Blackburn, who is wholly unsympathetic to what he takes to be Davidson’s position, calls the argument a “transcendental deduction of late twentieth-century American hegemony over the world of thought”, which depends not only on verificationism … but even worse on a particularly narrow conception of what verification of meaning involves, namely translation, or interpretation back into a home language. A more generous account would envisage the possibility of learning, and therefore of an expansion of the home language to embrace whatever new ways of thought may lie in our path. According to Blackburn, then, Davidson holds that we can understand a foreign language only if we can express whatever is expressible in the other language by means of our existing language. We could not expand our language or learn another.

---

Does Blackburn really think that Davidson believes this? The idea is patently absurd. It is an indisputable empirical fact that people sometimes do learn new languages, or new fragments of a language, so as to understand “new ways of thought”. One might learn the language of physics, for example, or music theory. Or, more radically, we can imagine people from a very primitive tribe coming to learn modern English or Japanese, and thereby coming to be able to say and understand things that would not have been possible by means of their original language. It is hard to believe that Davidson means to deny these kinds of obvious facts. In any case, it is clear that this argument, which is surely not Davidson’s, is unsound. Verificationism about meaning so extreme as to deny the possibility that we could come to understand concepts or beliefs other than those we presently have is clearly false.

I now turn to Rorty’s account of Davidson’s argument. Rorty takes Davidson to infer from the impossibility of discovering or recognizing an alternative conceptual scheme that there could not be such a thing. This reasoning would seem to be essentially what Davidson calls, in “On the Very Idea of a Conceptual Scheme”, “a very short line” against conceptual relativism:

Nothing, it may be said, could count as evidence that some form of activity could not be interpreted in our language that was not at the same time evidence that that form of activity was not speech behaviour. If this were right, we probably ought to hold that a form of activity that cannot be interpreted as language in our language is not speech behaviour. (VICS, 185-6)

This is a “short line”, I take it, in that involves a direct inference from the claim that some purported language is uninterpretable to the conclusion that it is not a language at all. This requires the assumption that if there is no intelligible criterion of identification for conceptual schemes, then the very idea of a conceptual scheme—i.e., the possibility of a workable criterion of identity for conceptual schemes—is also unintelligible.

However, as we’ve seen, Davidson recognizes that this “short line” is not particularly convincing; he writes that it depends on the assumption that “translatability into a familiar tongue [is] a criterion of languagehood” (VICS, 186). Since he concedes that this assumption is not obviously true, and needs to be supported by some argument. His discussion of the images and metaphors associated with conceptual relativism, which
we will explore in later chapters, is presumably intended to provide that support. So here, at least, Davidson appears to explicitly reject the verificationist “short line”, i.e., any direct inference from the “unrecognizability” of an incommensurable language to the conclusion that the purported scheme is not a language at all.

But it is at least debatable whether Davidson’s argument is, ultimately, verificationist. He seems to assume that to make sense of the idea that schemes “organize” a plurality of objects is to identify those objects: “whatever plurality we take experience [or reality] to consist in … we will have to individuate according to familiar principles” (VICS, 192). And he rejects the suggestion that conceptual schemes are true but untranslatable theories on the grounds that this provides no “test that a conceptual scheme is radically different from ours” (VICS, 195). This certainly seems to be a verificationist argument for his conclusion that we can’t make sense of the idea of an alternative scheme.

Moreover, even where Davidson appears to reject verificationism, his remarks are open to a different interpretation—or at least, they are not clear enough to decisively support either. On the one hand, he writes that if it is true that there could not be evidence of an alternative scheme, “we probably ought to hold” that there could not be such a thing: the impossibility of a test or criterion of identification for conceptual schemes is at least some reason to deny the intelligibility of the very idea. This suggests that Davidson does not clearly distinguish between criteria of identity and identification, and this would be most naturally explained on the assumption that his discussion rests on some form of verificationism. On the other hand, as we’ve seen, Davidson promises an argument for “translatability into a familiar tongue” as a “criterion of languagehood”. Presumably a verificationist would not see any need for such an argument. (If, as is surely true, we could not identify something as a language without thereby committing to the belief that is, at least in principle, translatable into our language.)

Hence it is unclear whether Davidson’s proposed “criterion of languagehood” is a criterion of identity or identification. His discussion of the verificationist “short line” does not settle the question of whether he thinks that the impossibility of recognizing an alternative scheme would imply that there could not be such a thing. And while I have suggested readings of his arguments against conceptual relativism that do not seem to rest
on verificationism, there is also textual evidence to suggest that these readings may be at odds with Davidson’s intentions. More precisely, it is ultimately unclear which thesis Davidson hopes to see “emerge as the conclusion of an argument” (VICS, 186):

(1) “Nothing … could count as evidence that some form of activity could not be interpreted in our language that was not at the same time evidence that that form of activity was not speech behaviour”; or,

(2) “A form of activity that cannot be interpreted as language in our language is not speech behaviour”.

According to Rorty, Davidson’s discussion is supposed to establish (1). Given verificationist assumptions, this would in turn establish (2). On the reading I favor, the discussion is supposed to establish (2), independently of any verificationist assumptions. (Of course, if (2) is true, so is (1). If the very idea of a language—i.e. the criteria of identity for languages—implies or depends on the concept of translation into our language, there could not be evidence that something is a language and yet is not translatable into our language. There could not be evidence for a hypothesis that is not even intelligible or coherent.)

II

On the verificationist reading, Davidson’s attack on conceptual relativism simply begs the question. Obviously the idea of an alternative conceptual scheme would be unintelligible if that were simply the idea of an “untranslatable” language that we can translate into our language. But, just as obviously, this is not the idea that the conceptual relativist claims to understand: he suggests that there are ways of conceiving of language that do not imply or depend on the “recognizability” of language by means of our language. It is this suggestion that needs to be evaluated, but an argument that simply presupposes the recognizability of all languages or schemes cannot do so.

In my view, Rorty’s reading is at odds with some important features of Davidson’s discussion. But my main concern will be to argue that, on Rorty’s reading, the argument against the intelligibility of the very idea of a conceptual scheme is unsound: the argument, on Rorty’s reading, is at odds with what I think the conclusion of
the argument *should* be, if the argument is of any interest. Let us consider Rorty’s reading in more detail. Rorty asks us to imagine that we encounter “humanoid organisms making sounds of great variety at one another, with what appear to be various effects on the interlocutor’s behavior”:

But suppose that repeated attempts to correlate these sounds with the organisms’ environment and behavior fail. What should we say? One suggestion might be that the analytic hypotheses we are using in our tentative translation schemes use concepts that we do not share with the natives … But could there be a way of deciding between this suggestion and the possibility that the organisms’ sounds are just sounds? Once we imagine different ways of carving up the world, nothing could stop us from attributing ‘untranslatable languages’ to anything that emits a variety of signals.  

We would never be able to decide that these sounds are really a language or, rather, “are just sounds”. And since this could not be decided, Rorty thinks, we would have reason to conclude that “the purported notion of an untranslatable language is as fanciful as that of an invisible color”.  

There are several *prima facie* difficulties with Rorty’s reasoning. First, this conclusion is inconsistent with Davidson’s own, that “we cannot intelligibly say that schemes are different, neither can we say that they are one” (VICS, 198). If some notion has no intelligible content, it cannot be “fanciful”. And indeed, if Davidson’s conclusion were merely that the idea of an alternative scheme is fanciful—i.e., silly or wildly improbable—this would not bear on the thesis that alternative schemes are logically possible. But the conceptual relativist needs only to defend the logical possibility of alternative conceptual schemes; this would be enough to suggest that the idea of an alternative scheme is intelligible. Of course, the vast majority of logical possibilities are “fanciful”, but they are nevertheless genuine (and intelligible) logical possibilities. Hence Rorty’s conclusion, as expressed, has no bearing on the possibility of an alternative scheme or on the legitimacy of the scheme/content distinction.

---

63 “Well Lost”, 653.
64 Ibid.
One might object to this that it is “very nearly, if not completely empty” to suggest that alternative conceptual schemes may exist, unbeknownst to us:

It is not enough to claim that there may be alternative schemes; to take conceptual relativism seriously we need to be presented with some reason to believe that alternative schemes exist. Unless a persuasive reason can be cited … the dialogue degenerates into an effort to shift the burden of proof to one’s adversary.  

However, Davidson’s thesis is not that there are no alternative conceptual schemes, or that we have no reason to believe that there are such things. Indeed, that thesis is of little philosophical interest, although perhaps, if it is true, it is interesting in other ways. (I suppose it would be interesting to discover that the language of the ancient Aztecs, say, was in fact perfectly interpretable by means of English.)

Rather, Davidson claims that the very idea of an alternative conceptual scheme makes no sense. And surely, if we are aiming to refute that thesis, it is enough to tell some seemingly coherent story, however improbable, about speakers whose language we aren’t able to interpret or translate into our language. That would, at least, constitute strong evidence against the claim that the idea of such a language is unintelligible. There is no need, in addition, to provide some reason for thinking that in fact there are such languages.

Second, it is not obvious that “once we imagine different ways of carving up the world, nothing could stop us from attributing ‘untranslatable languages’ to anything that emits a variety of signals”. Rorty seems to assume here that the only possible justification for attributing language (or, by extension, communication and thought) is translation or interpretation of the “signals” in question. If we suspend this condition on attributions of language, there is no more reason to suppose that the creatures in his thought experiment are speaking a language than to suppose that, say, lawnmowers are speaking a language.

This suggests a reductio of the assumption that there could be untranslatable languages, i.e. “different ways of carving up the world”: it is absurd to suppose that lawnmowers are speakers of an untranslatable language. But if we are prepared to suppose that anything else that “emits a variety of [uninterpretable] signals” is a speaker,

---

65 Marc Joseph, Donald Davidson (Stocksfield: Acumen, 2004), 180.
we should allow that lawnmowers are also speakers. Therefore, it is absurd to suppose
that anything that cannot be interpreted as speech behavior by means of our language is,
in fact, speech behavior.

However, we would have reasons independent of translation to attribute language
to “humanoid organisms making sounds of great variety at one another, with what appear
to be various effects on the interlocutor’s behavior”, but not to countless other things that
merely emit a variety of signals. Living creatures with sense organs, brains and bodies
highly similar to our own are, surely, much more plausibly characterized as speakers and
thinkers than inanimate objects, snails or amoebas. Moreover, if these creatures are
“making sounds of great variety at one another, with what appear to be various effects on
the interlocutor’s behavior”, the variety and complexity of their behavior further supports
the hypothesis that they are in fact speaking a language.

Thus, if Rorty’s reasoning is what it seems to be, it is invalid. Our common sense
and scientific knowledge of the many important differences between various kinds of
things that “emit a variety of signals” provides, at least, a far better reason for attributing
language to some of these kinds of things than to others. Since the reductio depends on
the false assumption that “anything that emits a variety of signals” is just as likely to be
speaking a language as anything else of that kind, it is unsound. To be sure, common
sense and scientific knowledge of the kind indicated above does not suggest a “criterion
of languagehood” independent of translation. (Or at least, independent of the logical
possibility of translation.) But it does provide us with criteria for making judgements as
to the likelihood that something is a language user, whatever exactly may be constitutive
of language.

Finally, Rorty’s claim that “repeated attempts” to interpret the behavior of the
creatures have failed or will fail is not sufficient to warrant the conclusion he draws. In
the situation he describes, it would at least be possible for us to imagine that, although no
successful translation has yet been discovered, there is such a translation. Perhaps we
might imagine that it will be discovered only in a thousand years, after a great deal of
scientific progress or cultural change. The mere fact that all attempts at translation, so
far, have failed, is not sufficient to establish that the purported language of the aliens is
“untranslatable”. (Except, perhaps, in some weak sense: it is untranslatable by any of the methods that have so far been applied to it.)

In any case, while it is no doubt true that given a sufficiently great number of failed attempts at translation, it would begin to seem unlikely that the aliens are speaking a language, this does not bear on the issue of intelligibility. (And again, if it seems increasingly unlikely that they are speaking a language, we must at least understand the hypothesis that, in fact, they are.) What Davidson presumably means to discuss is rather, as he says, “a language we could not translate at all” (VICS, 192). At any rate, this is what he should be discussing. But this is not the same as a language that we have not, yet, been able to translate.

III

Rorty’s interpretation of Davidson’s argument must therefore be qualified in a number of important ways if it is to bear on the question of whether the notion of an alternative conceptual scheme is intelligible:
(i) The conclusion of the verificationist reading of Davidson’s argument should be that the idea of an untranslatable language (an alternative conceptual scheme) is unintelligible, not that it is “fanciful”, bizarre, etc.
(ii) The argument should not depend essentially on the false assumption that nothing short of a translation of some language into our own could serve to distinguish the behavior of speakers of that language from that of “anything [else] that emits a variety of signals”.
(iii) The argument should not rest merely on the premise that “repeated efforts” over some fairly long period of time have failed to produce a translation of the putative language. Rather, what is in question is whether we can make sense of the idea of “a language we could not translate at all”—that is, regardless of how hard we try, and for how long.

In light of these qualifications, Rorty’s reasoning is quite unconvincing. At most, it establishes the uncontroversial claim that in the kind of situation he imagines, we would have little reason, or perhaps no reason at all, to suppose that these alien beings
were speakers of a language. This is consistent with the conceptual relativist’s claim that they might nevertheless be speakers of a language.

I have tried to establish two main conclusions in this chapter. First, the mere fact that some purported language might be untranslatable into ours at some time provides no good reason for the claim that the purported language is not a language at all. It seems perfectly intelligible to hold that something is a language despite being “unrecognizable” as a language, in the sense that we, with our actual powers and knowledge, are not able to recognize it as a language. Second, the more extreme case of a purported language that is “unrecognizable” as a language, in the sense that we could not, in principle, translate it into our language, also seems to be intelligible. As Davidson concedes, it is not obvious, at least, that the possibility of translation into our language is a necessary condition of language-hood. Some argument is needed for this claim.

Therefore, even if some purported language is unrecognizable or untranslatable in the strong sense suggested just now, a direct inference to the conclusion that the purported language is not a language at all appears to be invalid. In the next chapter, I’ll consider some interpretations of Davidson’s argument that, unlike the verificationist “short line” considered here, hinge on verificationist premises that concern specific aspects of various formulations of the scheme idea. These interpretations thus involve an indirect inference from the impossibility of translation to Davidson’s conclusion that the very idea of a conceptual scheme is unintelligible. As we’ll see, these indirect inferences also appear to be invalid.
CHAPTER 3
ORGANIZATION, PART 1: EXTENSIONAL DIFFERENCES

Davidson’s arguments in “On the Very Idea of a Conceptual Scheme” focus on two general characterizations of conceptual schemes. In this chapter and the next, I’ll explore his treatment of the metaphor of organization. On this kind of view, conceptual schemes are to be understood as incommensurably different systems of organizing principles. First, however, I’ll discuss Davidson’s master argument against the intelligibility of the idea of a conceptual scheme, and the role of his discussion of the metaphor of linguistic organization in that argument.

We noted earlier that Davidson holds that conceptual schemes (if there are such) are identical or at least very closely associated with sets of inter-translatable languages. Essentially, his view is that conceptual schemes differ if and only if the languages with which they are associated are not translatable into another. (In fact this formulation will require some important qualifications, as we’ll see, but we can ignore these for the moment.) This is the first important step in his master argument, and it is worth briefly considering why this might be an acceptable criterion of identity for conceptual schemes.

There are some compelling prima facie grounds for accepting Davidson’s criterion. Pre-analytically, it is natural to think that to translate or interpret a speaker’s language is to uncover his concepts, beliefs or view of the world. For example, if we can correctly translate a speaker’s utterance of “Schnee ist weiss” by our sentence “Snow is white”, then presumably we share with that speaker the concepts of snow and whiteness. By contrast, if some language were in principle untranslatable into our own, one might well conclude that the system of concepts associated with that language was incommensurable with our own. Hence translatability seems to be a fair criterion for sameness or difference of conceptual scheme as concerns thinkers whose use of concepts is mediated by language. Moreover, as we’ll see, the association of language and conceptual scheme is a common theme in the accounts of conceptual schemes that Davidson discusses—e.g., in Quine’s philosophy.
But could there perhaps be a kind of conceptual thought without language? If so, there might be conceptual schemes for which the criterion of translatability would be inappropriate. Davidson would not grant the possibility of conceptual thought without language, because he holds that conceptual thought is essentially inter-subjective: the contents of a speaker’s thoughts are (partly) determined by his relations to other speakers. We’ll explore this key feature of Davidson’s philosophy in more detail later. In what follows, I will assume Davidson’s criterion. If one has concepts, these are, or at least could be, represented in speech. Inter-translatability is thus the condition of identity for conceptual schemes (if such there be).

I.

Schemes differ, then, where the language associated with one cannot be translated into that associated with another. But what exactly makes a language, or a set of languages, a conceptual scheme? Davidson explores the notion of a conceptual scheme through the dualism of schemes and the content supposed to be common to divergent schemes. This dualism concerns, on the one hand, a language, which encodes a conceptual scheme, a perspective, theory or world view, and, on the other hand, an entity that the language represents or interprets, which other schemes may represent or interpret differently. So a language is related in some specified manner to a specified entity. The relation of language qua conceptual scheme to this entity is supposed to suggest an explanation of how differences in scheme lead to failure of translatability: schemes differ when languages are related to the content in incommensurally different ways.

How might such failures of translation arise? According to Davidson, there are essentially four main ways in which languages might be (or be associated with) conceptual schemes:

The images and metaphors fall into two main groups: conceptual schemes (languages) either organize something, or they fit it: The first group contains also systematize, divide up (the stream of experience); further examples of the second group are predict, account for, face (the tribunal of experience). As for the entities that get organized, or which schemes must fit, I think we may again detect two
main ideas: either it is reality (the universe, the world, nature), or it is experience (the passing show, surface irritations, sensory promptings, sense-data, the given). (VICS, 191)

Davidson takes these “images and metaphors” to indicate, in different ways, the dualism of scheme and content, which he says is a “dogma of empiricism” (VICS, 189). The dualism is “the foundation of an empiricism shorn of the untenable dogmas of the analytic/synthetic distinction and reductionism” (VICS, 189). For even if we reject these other “dogmas”, we may still hold that there is a distinction between language or theory, on the one hand, and sensory experience or reality, between “organizing system and something waiting to be organized” (VICS, 189). (As we noted earlier, this is misleading, given Davidson’s other remarks about the dualism of scheme and content: as he describes it, it is not peculiar to empiricism.)

According to Davidson, then, there are essentially four ways in which languages might be conceptual schemes. Languages *qua* conceptual schemes may be said to (i) organize reality, (ii) organize experience, (iii) fit reality or (iv) fit experience. If any of these suggestions can be elaborated in a way that does not imply or depend on translatability into our language, the idea of a difference in conceptual scheme—and so, by extension, the very idea of a conceptual scheme—may prove intelligible. To conceive of an alternative scheme, we need only imagine something that stands in the appropriate relation to reality or experience, and yet is not translatable into our language. Davidson then proceeds to argue that all of proposals (i)-(iv) turn out to imply or presuppose the possibility of translation into our language. More precisely, Davidson argues that we cannot make sense of the idea of an alternative conceptual scheme, i.e., an untranslatable language, because translatability into *our* language is a “criterion of languagehood” (VICS, 186). If all other languages are translatable into our language, they must all be inter-translatable.

The immediate conclusion is thus that we cannot make sense of the idea that there are *many* conceptual schemes; there is at most one. However, Davidson holds that the very idea of a conceptual scheme implies the possibility of others: “if we cannot intelligibly say that schemes are different, neither can we intelligibly say that they are one” (VICS, 198). Without at least the possibility of an alternative to our own conceptual
scheme, i.e., an alternative way of organizing or fitting experience or reality, there is no way to distinguish between scheme and content: no way to distinguish between the subjective and objective features of our experience or view of the world. If any language or system of concepts must be intelligible to us, by way of translation into our language, there is no distinctly subjective contribution to our experience or world view: no subjective form or framework, imposed on reality or experience, that could be radically different. The ultimate conclusion of the argument is therefore that all talk of conceptual schemes is empty.

So, in outline, Davidson’s master argument is roughly as follows:
(1) The idea of a conceptual scheme is intelligible only if the idea of a difference in conceptual scheme is intelligible.
(2) The idea of a difference in conceptual scheme is intelligible only if the idea of a language that cannot be translated into our language is intelligible.
(3) The idea of a language that cannot be translated into our language is intelligible only if there is some conception of language that does not imply or depend on the concept of translatability into our own language.
(4) There is no such conception of language.
Therefore,
(5) The idea of a conceptual scheme is not intelligible.

The most contentious step in the argument is, I think, the claim that there is no conception of language that does not involve the possibility of translation into our own language, i.e., that “translatability into a familiar tongue” is “a criterion of languagehood” (VICS, 186). Davidson acknowledges that this “lacks the appeal of self-evidence”, and so “should emerge as the conclusion of an argument” (VICS, 186). The argument for this key premise consists in a range of subordinate arguments directed against the intelligibility of metaphors (i)-(iv), which are supposed to indicate ways in which languages qua conceptual schemes might fail to be inter-translatable, despite being related in some special way to experience or reality.

These subordinate arguments tend to substantiate Davidson’s observation that the dominant metaphor of conceptual relativism, that of differing points of view, seems to betray an underlying paradox. Different points of view make sense, but
only if there is a common co-ordinate system on which to plot them; yet the existence of a common system belies the claim of dramatic incomparability. (VICS, 184)

Thus, “it is essential” to the scheme idea “that there be something neutral and common that lies outside all schemes” (VICS, 190)—reality, the world, or instead experience, sense data, etc. And yet, it is also essential that this neutral and common thing could not, even in principle, serve to establish a translation between the languages associated with divergent schemes. In the abstract, this line of thought does seem deeply paradoxical, as Davidson emphasizes. If A and B are systematically related to C, surely A and B are also systematically related to each other. If A and B are languages, we might suppose that their relations to a common object would enable us to devise a translation of either language into the other.

Of course, there is much more to be said about this line of argument. As it stands, it is at best suggestive. The mere fact that there are systematic relations between each of two schemes and some third entity, and so, surely, between each scheme and the other, does not directly establish that the two schemes (or the languages with which they are associated) are inter-translatable. Translation establishes a particular kind of systematic relation between two languages, and it may be that the relations between divergent schemes are too weak, or of the wrong kind, to yield a translation of one into the other.

In developing this line of objection, therefore, Davidson needs to provide some reason for thinking that the relations between schemes or languages and some third entity are sufficient for translation. A number of such reasons emerge from an examination of the details of various accounts of conceptual schemes as languages that organize or fit some “neutral and common” entity; the reasons concern the real features of language involved in the metaphors of organizing and fitting, which tie language to the objective world, and hence to other languages.

For now, I want to emphasize that Davidson’s attack on the various metaphors of conceptual relativism can be seen as an illustration, by reference to some typical and influential accounts of conceptual schemes, of what he takes to be a basic defect in the very idea. He hopes to show that whatever exactly the relation between scheme and content may be, and whatever exactly the entity to be contrasted with language, the very
idea of a conceptual scheme involves two conditions that cannot both be met, because they are inconsistent.

The central phase of the master argument thus plays two complementary roles in Davidson’s attack on the scheme idea. On the one hand, this phase of the argument is an instance of *modus tollens*. If languages (or sets of inter-translatable languages) are associated with conceptual schemes, then at least one of metaphors (i)-(iv) must allow for the possibility of a language so related to a neutral “content” that is not translatable into our language. But, according to Davidson, none of these metaphors allows for this possibility. Therefore, languages cannot be intelligibly identified or associated with conceptual schemes. On the other hand, this same phase of the argument can be taken to indicate a reason why these various accounts of conceptual relativism are paradoxical, and in fact incoherent, by revealing features of language and conceptual thought that depend essentially on the possibility of translation and interpretation. We’ll explore this suggestion in more detail in later chapters.

II

Let us now consider in more detail Davidson’s argument against the proposal that languages or conceptual schemes can be characterized as systems of organizing principles, in a way that does not imply or presuppose the possibility of translation into our language. The first step in the argument is an appeal to the content of the metaphor of organization. Davidson writes that “we cannot attach a clear meaning to the notion of organizing a single object (the world, nature, etc.) unless that object is understood to contain or consist in other objects” (VICS, 192). One might organize the shirts and shoes in a closet, but not the closet itself. (Unless, of course, the closet is “understood to contain or consist in other objects”—e.g., a floor, walls and ceiling, or wooden parts and plaster parts.)

Thus, if we are to suppose that languages or schemes organize reality (or the world, nature, etc.), reality must be understood to consist in a plurality of objects. But in that case “we will have to individuate” these objects “according to familiar principles” (VICS, 192). These must be objects we can identify and describe. (Davidson mentions
such things as “knives and forks, railroads and mountains, cabbages and kingdoms” (VICS, 192).) This conclusion undermines the claim of incommensurability. If some other language organizes objects that we can identify and describe according to the “familiar principles” of our own language, that language is one that we can, presumably, come to understand. We can “plot” the relations between each language and recognize the concepts, categories or organizing principles of the other language by means of a “common co-ordinate system”, the plurality of objects organized by both languages.

Notice that Davidson’s position is not that, because we can individuate the objects in question according to familiar principles, the organizing principles of other languages must also be familiar—it is not that our concepts are the only ones possible. Of course there are, or at least there could be, languages that organize familiar objects in ways that we would find “strange or novel” (VICS, 184). Rather, Davidson’s position seems to be that the objects organized by a language constrain the organizing principles applicable to those objects, and hence it would be possible for us to come to understand these strange or novel forms of conceptual organization by means of our own language or concepts.

If kingdoms, say, are the objects to be organized, they can be organized in some ways but not in others—by geographic position or ruling family, say, but not by musical key or shoe size. The organizing principles applicable to kingdoms may be highly diverse, but they can’t be so diverse that speakers whose language organizes kingdoms in one way would be unable to understand the utterances of speakers whose language organizes them in another way. Since we can identify and describe the objects organized by purportedly alternative schemes, we can come to understand those other schemes by means of the objects they organize. But then they are not alternative schemes after all. Davidson concludes that “the image of organizing the closet of nature” (VICS, 192) fails to provide a conception of language independent of translatability into our own language.

The same reasoning applies if the entity organized is, instead, experience, sense data or sensory promptings. In that case, the plurality of objects common to supposedly different and incommensurable schemes will presumably consist of “events like losing a button or stubbing a toe, having a sensation of warmth or hearing an oboe” (VICS, 192). Whatever exactly the objects organized may be, they must be objects we can individuate according to familiar principles. Hence, the conclusion is, again, that “any language that
organizes such entities must be a language very like our own” (VICS, 192). Davidson notes that this formulation of the scheme/content relation is open to an additional objection, in that it is unclear how we could conceive of a language “that organizes only experiences, sensations, surface irritations or sense-data” (VICS, 192).

Although Davidson does not dwell on this objection, it is easy to see a number of ways in which the idea of such a language is implausible or even unintelligible. Second, it is hard to imagine how such a language could be learned, or how it could serve as a means of communication. Without reference to a shared reality, how could anyone learn the meaning of words and sentences? Seemingly, to learn this kind of language would require that the student be able to perceive the teacher’s sensations. Otherwise, he will not be able to grasp the reference of the teacher’s words, but of course this is impossible. Davidson mentions another, less important objection. Our own language plainly organizes far more than experiences, sensations and the like. Something that organizes only experiences, it would seem, is only a fragment of a language.

III

We’ve seen that Davidson’s main line of argument against accounts of the scheme idea based in the metaphor of organization is intended to apply whether the entity organized is reality or experience (or even, for that matter, some third entity). Let us consider this main line of argument more closely. Davidson’s first premise seems undeniable. Only pluralities can be organized, arranged, systematized or divided up. Unless the thing being organized has parts or contents, there is nothing to organize. (Of course, the parts or contents need not be actually separate and distinct. They need only be separable or distinguishable—as are (say) the left and right sides of a wall.)

The idea of organizing reality or experience, then, implies that the entities to be organized must “contain or consist in” a plurality of parts or contents or regions. What is less clear is why we should accept Davidson’s claim that whatever plurality of objects is organized by supposedly divergent schemes must be one that we can individuate according to familiar principles. This formulation of the scheme idea does seem to require that there is some plurality of objects that divergent schemes organize, but it
doesn’t follow from this that we can identify or describe them—or, for that matter, that speakers of any language are able to do this.

An argument for this premise can be drawn from Davidson’s suggestion that the notion of conceptual or linguistic organization concerns “the referential apparatus of language—predicates, quantifiers, variables and singular terms” (VICS, 193). Reference, predication and quantification serve to classify, categorize or divide up reality into individuals, parts, regions and kinds. Moreover, it is hard to imagine how else the notion of linguistic or conceptual organization could be understood. Languages can’t physically separate shoes from socks, for example, or arrange books on a shelf. So what is at issue here is essentially the use of predicates and other referential devices.

The idea of an alternative conceptual scheme, then, is the idea of a system of predicates that applies to (organizes) the same objects as our own system of predicates, but is not translatable into our language. But then we, with our predicates, can individuate the objects to which the alien predicates apply. Of course, we presumably do not individuate the same *collections* of objects—perhaps we classify shoes by color or shape whereas the aliens classify those same things by sensory properties we can’t detect. In any case, this scenario would seem to require that the plurality of objects organized by our scheme and the other consists in objects we are able to identify. Davidson’s point would seem to be simply that, since linguistic or conceptual organization involves predication, it requires a discriminatory ability. If we understand or make use of the predicate “dog” (or the concept of a dog), we must be able to distinguish dogs from other things. Thus, if our language and some other both organize dogs (say), then speakers of both languages must be able to individuate dogs, even if, as may be, they classify or conceive of dogs in very different ways.

Some commentators take Davidson’s suggestion here to be, instead, that the only kind of “referential apparatus” we can intelligibly imagine or discuss is one “much like our own”. On this reading, Davidson holds that the conceptual relativist must identify or describe the organizing principles of other schemes. To provide such a description requires a translation of at least some of the predicates of alien languages into the translator’s own language; therefore, the idea of a completely untranslatable language is not intelligible. The argument thus appears to rest on some form of verificationism.
Unlike Rorty’s interpretation of the argument, this interpretation does not involve a direct inference from the impossibility of translation to the conclusion that purportedly untranslatable languages are not languages at all. Instead, the argument here depends on considerations specific to the metaphor of organization: given the nature of this characterization of alternative schemes, they turn out to be identical with our own.

Thus, David Henderson, for instance, agreeing with what he takes to be Davidson’s position, writes:

If something is to be a language that organizes objects, it must have predicates picking out classes of objects. But, we can only make sense of a language doing this, when we can translate at least many of its predicates. So, Davidson concludes that the idea of a language that organizes objects is the idea of a language that is at least partially translatable into our own. It must be one for which we can determine the extension of many of its predicates.\(^66\)

Henderson thinks that we cannot “make sense” of the idea that something is a language or system of organizing principles unless we can translate much of it into our language, and this means we cannot make sense of the idea of a radically alien system of organizing principles.

I find this line of argument unconvincing, for reasons essentially the same as those advanced in the last chapter. The proposition that John is organizing something seems to me perfectly intelligible as it stands: it seems to me that I can make sense of that proposition without knowing, assuming or imagining that he is organizing books, say, rather than shoes or pianos, or that he is organizing them alphabetically rather than chronologically, etc. Similarly, to make sense of the claim that something is “a language that organizes objects” is just to understand the proposition that that thing “organizes” objects: that it has predicates and singular terms that pick out objects, quantifiers that range over objects, etc. Why is it necessary, in order to understand that proposition, that one know or imagine, in addition, which objects it organizes, or how exactly it organizes them? Surely we could imagine that something is a word or sentence, true of certain

---

objects or under certain circumstances, without presently being able to determine which ones.

Much the same kind of objection applies to Henderson’s more general formulation of what he takes to be Davidson’s reasoning:

To attribute disagreement, we must be in a position to identify the content of the relevant beliefs. But, these putative beliefs can only be assigned content (indeed, only have the content they have) by being “anchored” or located in a patterned network of associated beliefs… Thus, differences must be localized in that disagreements can only be identified if the particular belief (and the concept) at issue is fairly directly anchored by its association with many related beliefs (having to do with related concepts) on which there is agreement.\footnote{“After Davidson”, 189.}

All that clearly follows from the assumption that beliefs can be ascribed or identified only in relation to some larger pattern of beliefs is that identifiable conceptual or theoretical differences must be localized. It does not follow that unidentifiable differences between languages or theories or systems of concepts could not be global.

Similarly, one might allow that the identity of a belief is determined by its place some larger “network of associated beliefs”, but the strongest conclusion we might draw from this, however, is that we would not be able to identify any beliefs belonging to a totally alien pattern. The conceptual relativist might without inconsistency concede this much—indeed, it would seem to an essential part of his position—and hold that there nevertheless are, or could be, radically alien systems of belief.

Alternatively, Davidson’s reasoning might be that since, by hypothesis, all languages organize the same objects, they all organize whatever it is that our own language organizes. Our language organizes things like cabbages and kingdoms, since it is these familiar kinds of objects over which our quantifiers range and of which our terms are (sometimes) true. Therefore, all languages organize or refer to some or all of the objects that figure in the semantics of our own language. The ontology of English thus turns out to be “something neutral and common that lies outside all schemes” (VICS, 190), and we can appeal to that ontology to interpret any purportedly alien conceptual scheme.
This argument is too quick, however, as the conceptual relativist is likely to hold that the ontology of English is itself an artifact of some prior organization of other, unfamiliar objects, such as sense data or sensory promptings. The claim that our own language organizes or refers to the familiar objects Davidson mentions thus begs the question. For it might be that we could know the real or ultimate referents of our own language only if we could do precisely what the conceptual relativist claims is impossible—namely, set aside our own language or scheme and compare it with the “neutral and common” reality it organizes.

IV

Davidson often appears to endorse the reasoning I have been criticizing. Thus, while he concedes that “a language may contain simple predicates whose extensions are matched by no simple predicates, or none at all, in some other language” (VICS, 192), and that translation may fail in such cases, he seems to argue, in a key passage, that this kind of scenario is intelligible only if we can identify and understand the conceptual or linguistic differences that prevent translation:

What enables us to make this point [i.e., that the extensions of predicates in one language have no counterparts in another is impossible] in particular cases is an ontology common to the two languages, with concepts that individuate the same objects. We can be clear enough about breakdowns in translation when they are local enough, for a background of generally successful translation provides what is needed to make the failures intelligible. But we were after larger game: we wanted to make sense of there being a language we could not translate at all. (VICS, 192)

Davidson’s argument in this passage could be taken to be a version of Henderson’s, cast in linguistic rather than mental terms: the claim that one language is not translatable into another is intelligible (i.e., “clear enough”) only if it is possible to identify and understand the differences in extension between the two languages (i.e., only if it is possible to “make the failures intelligible”). However, this is possible only if we can compare both languages with “a common ontology”, a single plurality of objects individuated in different ways. Of course, if we can make this kind of comparison, we
must possess a third, more comprehensive language, which incorporates both the first two languages as fragments. Therefore, it is intelligible to claim that one language is not translatable into another only if both languages are fragments of a single language and system of concepts; moreover, this must be a language or system that we ourselves possess, or could come to possess.

But why should we accept the premise that the notion of a failure of translation is intelligible only when we can “make the failures intelligible”? The argument might be thought to rest on a suppressed verificationist premise, to the effect that the intelligibility of failures of translation is somehow essentially connected to our ability to “make the failures intelligible”. However this would, again, beg the question against the conceptual relativist, by tacitly assuming the controversial thesis that Davidson originally promised to support by some argument.

Alternatively, Davidson’s argument might be an induction over “particular cases”: whenever we are able to make sense of this kind of failure of translation, the languages in question can (eventually?) be seen as fragments of a more comprehensive language, in relation to which the failures are local and fairly shallow. Needless to say, this is not a particularly strong argument against the very idea of a more radical, global failure of translation. Perhaps this induction supports the much weaker conclusion that such a situation is very unlikely; but this is not to say that the very idea is somehow unintelligible.

On either reading, the argument fails because the conceptual relativist claims merely to imagine a certain kind of situation, a kind of situation, moreover, which seems readily intelligible. This is not to claim that we would be able to identify or describe particular instances of this kind of situation. If Davidson’s argument is directed merely at the claim that we would be able to identify or describe particular instances of this kind of situation, it establishes nothing about the intelligibility of the bare claim that there could be situations of that kind. Nor, arguably, does it even engage with the conceptual relativist’s position, since the claim that we could identify or describe instances of that kind is not essential to that position; indeed, the two are probably inconsistent.

However, the key passage cited above is open to a different and better reading, which is directed, as it should be, against the very idea of a conceptual scheme (on the
present account of conceptual schemes). First, notice that in this passage Davidson does not claim that languages related in this way must be associated with the same concepts. If it were true, this would be sufficient to refute the suggestion that languages that “organize” a single collection of objects might be associated with incommensurably different concepts; but this is clearly not true. Languages that serve to classify or individuate a single collection of objects may be associated with very different systems of concepts.

As Davidson says elsewhere, “some people have conceptual resources not available to everyone”:

Biologists, aeronautical engineers, solid state physicists, musicologists, cartographers, molecular biologists, selenographers, and psychoanalysts all command vocabularies and theories many of us do not. In our more restricted way, we common types also have our specialties: our own list of proper names, with their uniquely contextualized references, our private endearments and verbal twists, our own mispronunciations and malapropisms.68

There are many familiar cases in which different systems of concepts “individuate the same objects” (VICS, 192), and this does not suggest a difference in conceptual scheme. The languages of biology and psychoanalysis, for example, provide us with different ways of talking about human beings and human behavior, but of course the concepts associated with those languages are quite different. Presumably Davidson’s rejection of the scheme idea is not a denial of these familiar kinds of conceptual differences.

Instead, Davidson’s point here seems to be that the kind of failure of translation under discussion requires a common ontology, a common plurality of objects, in relation to which concepts differ. The existence of a common ontology suggests that interpretation between speakers of these different languages (or language fragments) is, in principle, possible. That is to say, the particular cases that we can “make intelligible” by reference to a “background of generally successful translation” reveal the features of the kind of situation described by the conceptual relativist. Where translation fails because of a difference in organizing principles, there is some shared ontology, and some

discrepancy—perhaps an enormous discrepancy—between the extensions of the predicates of two languages. So, on this reading, unlike Henderson’s, Davidson’s argument directly addresses the claim that a global translation failure of this kind is possible, regardless of our ability to identify instances. Let us consider this line of argument in more detail.

On this interpretation, Davidson’s view is that in this kind of situation each of the two languages is, in principle, translatable into the other (even if, as may be, this requires that both be translated into some third language). Such translation, or interpretation, at least, must be possible given merely the assumption that both languages organize, individuate or refer to the same ontology. The relations between each language and those common objects are transitive. If the Martian word “kibbage” refers to certain cabbages and certain kingdoms, it also refers to the referents of certain English words, and there is a meta-language by means of which to contrast the semantics of Martian and English. It does not follow from this that we ourselves must always be able to learn or devise a language by means of which a “generally successful translation is possible”. The conclusion is simply that such a language is possible, and therefore the original languages are not associated with distinct conceptual schemes.

So perhaps Davidson’s claim that we can “individuate according to familiar principles” the objects of purportedly different schemes is less a premise than an alternative statement of his conclusion. The main point is that differences between organizing principles cannot be “so extreme but that the changes and contrasts can be explained and described using the equipment of a single language” (VICS, 184). Strictly speaking, then, he should not say that we must be able to individuate the objects of purportedly different schemes according to familiar principles. We might not be able to individuate them at all, or only according to new and unfamiliar principles. Rather, the point is that some single language or system of concepts and principles has the resources to “make the failures intelligible”, and that this interpretation would (of course) proceed according to principles that can be seen to underlie both languages.

To put this point another way, suppose that in some such case of translation failure there could not be any such third language—i.e., that there could not be any way of representing the extensions of expressions in either of the first two languages as
identical with all, some or none of the extensions of expressions in the other. It is hard to imagine how this could be impossible in principle, unless the languages in question do not in fact organize the same collection of objects. Given the assumption merely that the objects in question can be named (in some language), it should be possible in principle to identify the extensions of any predicate in one language with some, none or all of the objects in the extensions of predicates in the other. As we’ve seen, he does not commit to the claim that inter-translatable languages must be associated with the same concepts (i.e., that they must contain predicates with identical extensions).

As suggested earlier, then, Davidson’s discussion can be seen as an instance his more general claim that conceptual relativism is paradoxical. We can make sense of the idea that languages, in their organizing capacity, are like points of view or perspectives, but only if they are supposed to differ in relation to a common ontology. But this forces the conclusion that there is “a common co-ordinate system” (VICS, 184) against which the conceptual differences could, in principle, be understood and described by means of a single language.

V

As a further illustration and defense of what I take to be Davidson’s position, consider P.M.S. Hacker’s suggestion that we can at least make sense of the notion of a partial difference in conceptual schemes (qua organizing principles) by imagining divergent criteria for the application of concepts:

A linguistic community which employs different ranges of colour samples from those we do … will employ different criteria of sameness and difference of colour. To that extent, not only their concepts of colour, but also their concept of colour will be somewhat different from ours, even though it is still rightly deemed to be a concept of colour. In saying, in the native tongue, that a certain rose is \( N \)-coloured, they may indeed be saying something true, but not translatable into English, since ‘\( N \)’ is not equivalent to any colour predicate in English… [The native word] may be explained by reference to samples we should call ‘red’ and samples we should call ‘orange’, but samples we should call ‘orange’ are also used for a further colour
predicate which covers part of the orange and the whole of the yellow range. So there is no form of words in the native language for ‘red but not orange’. Hacker seems to presuppose an unusually stringent notion of translatability. There is not presently a single English word that translates the native word ‘N’, but there is a phrase that does: “red or orange”. And, of course, English speakers can of course coin a new word co-extensive with ‘N’, if they like: “rerange”, say. So a word for word translation of the native language into English is possible after all, unless we are to suppose that the addition of even one word to the existing vocabulary of English would constitute a change of language. For the same reason, nothing of interest follows from the fact that “there is no form of words in the native language for ‘red but not orange’”. There may presently be no such expression, but the natives could invent one by discriminating more finely between what they normally take to be shades of rerange.

Admittedly, there are more extreme cases of this kind, where it may be less plausible to speak of translation. Dorit Bar-On offers the following example:

_Mikabari_ is a word of Malagasi meaning, roughly, to perform a _kabary_: a special kind of formal speech given only on certain types of ceremonial occasions. _Mikabari_ stands for an element in the social background of the people of Madagascar, absent from that of English-speaking people. To convey the content of _mikabari_ in English, one would have to provide a very detailed explanation, possibly of book length, of certain activities and cultural institutions peculiar to the Malagasi speaker and unfamiliar to the English speaker. Even by the loosest standards, such an explanation would not qualify as a translation of the word. Henderson gives the example of Evans-Pritchard’s attempt to explain the meaning of the Zande word “mangu”:

*Mangu* (1) WITCHCRAFT SUBSTANCE: a material substance in the bodies of certain persons. It is discovered by autopsy of the dead and is supposed to be diagnosed by oracles in the living. (2) WITCHCRAFT: a supposed psychic

---

emanation from witchcraft substance, which is believed to cause injury to health and property.\textsuperscript{71}

We can "convey the content" of "$N$" in our language by means of a very brief explanation of meaning, and the natives can do the same for our phrase "red or orange", although this will require them to give a somewhat longer and more complicated explanation. But an explanation in English of the meaning of "mikabari" or "mangu" will be much longer and much more complicated, and may require English speakers to "reconstruct" a wide range of "theoretical" concepts on the basis of other concepts shared with speakers of Malagasi or Zande.

Perhaps, as Bar-On and Henderson hold, such accounts cannot always be plausibly described as "translations" of the alien expressions. According to Bar-On, this is because translation "a pairing of source discourse with conditions of truth formulated in the target language", while an explanation of the meaning of "mikabari" would involve, rather, "metalinguistic explanations of background (linguistic, socio-cultural, historical)".\textsuperscript{72} Similarly, Henderson distinguishes between what he calls "strict" and "reconstructive" translation. The Zande words for "house" or "house", say, may be amenable to "strict" translation into English. The meaning of "mangu", however, is related in only in some respects to what we call "witchcraft"; such words can be "[fully] explicated only in the holistic reconstruction of a substantial fabric of Zande belief—the Zande theory".\textsuperscript{73}

I am not sure that these are principled distinctions. Hacker’s account of what it meaning of ‘$N$’, for example, could be seen as a “metalinguistic explanation” of culture, history, etc., or a “reconstruction” of the native “theory”—albeit an explanation that is very brief and straightforward compared with that required to explain the meaning of “mikabari” or “mangu”. If the natives do not differ from us in too many other important respects, the relevant facts about the metalinguistic background can go without saying. But whether the metalinguistic “background” is taken for granted or, instead, has to be


\textsuperscript{72} "Conceptual Relativism", 160.

\textsuperscript{73} "After Davidson", 179.
made explicit, it has an essential role in explaining or coming to understand the meaning of utterances in another language (or, for that matter, in one’s own language).

Alternatively, a book length English account of a foreign culture, religion and history that conveys the meaning of “mikabari” or “mangu” could be seen as a translation into English of Malagasi or Zande expressions. Such an account would, after all, state the truth conditions of the claim that Jones is performing a *kabary*. This statement of truth conditions would perhaps be very long and complicated, and would depend on statements of the truth conditions of many more, related Malagasi utterances and extra-linguistic facts than would a translation of ‘N’. But it isn’t clear to me why the mere fact that some explanations of meaning are longer, more complex or mentally demanding than others should make the difference between translation and some other activity that produces an account in one language the meaning of expressions in another.

Ultimately, however, it doesn’t matter how exactly the concepts of *translation* or *language* are normally understood, or should be understood. What is really at issue here is the possibility of understanding or interpretation across different languages or systems of concepts. Translation failure was introduced as a criterion for differences in conceptual scheme only because it seemed to reflect a failure of understanding. On a sufficiently narrow characterization of language or translation, translation and understanding can of course come apart. If a narrow characterization of language or translation seems necessary or useful for some purposes, we can reformulate Davidson’s position without any essential dependence on these notions. The onus is on proponents of the scheme idea to provide some basis for distinguishing between languages and identifying cases of translation failure. Unless they can offer some such account, the thesis that there could be non-inter-translatable languages has no clear content.

So the view I defend, and that seems the most charitable reading of Davidson’s discussion, is that on any reasonable account of language or translation, there cannot be purely linguistic (conceptual) factors that would rule out the possibility of mutual understanding between speakers whose linguistic (conceptual) behavior is referentially related to the same plurality of objects. In making sense of another’s utterances, we may often need to introduce new terms into our vocabulary, but it seems unimportant whether we choose to say that this is merely a modification of our existing language or, instead,
the acquisition or invention of an altogether new language. The first description will be more natural in some cases, the second in others, but regardless of how exactly we describe this activity, linguistic or conceptual differences of the kind considered so far will not prevent mutual understanding.

Of course, there may be other, non-conceptual reasons why mutual understanding is difficult or impossible. Thus, Hacker writes that if the natives are all color-blind, “they may have a colour predicate ‘M’ that covers the range of red, green and grey”:

They will rightly say that poppies, grass and clouds are the same colour, namely, M. This we could translate as the claim that poppies, grass and clouds are either red, green or grey—which is indeed true. But our true assertion that poppies differ in colour from grass is not something that could be translated into their language because the meaning of the term ‘colour’ is partly determined by the criteria of identity and difference for colour.

This conclusion is far too strong. Any English speaker with normal eyesight will be able to translate English into the native language, by adding a few new terms to its vocabulary. (These should include terms that distinguish between what we, in English, might call “English-color” and “Native-color”, so as to express in the native language the proposition that poppies differ in color from grass.)

Although Hacker claims to have established that “the native conceptual scheme or system of representation … cannot (logically) be mapped (or mapped perfectly) on to our conceptual scheme, precisely because the grammar of colour differs between the two languages”, it appears that, on the contrary, it can quite easily be mapped (perfectly) on to ours. There is nothing about the natives’ language or concepts that prevents them from understanding certain English sentences; there is something wrong with their eyes. Differences in conceptual scheme, by contrast, are supposed to involve some special relation of incommensurability between languages or systems of concepts. It is this special relation that, supposedly, makes understanding impossible. Cases in which understanding is difficult or impossible for some other reason are irrelevant.

74 “Davidson’s Idea”, 304.
75 Ibid., 305.
This point deserves emphasis because it is often overlooked. Alex Byrne offers the following argument against what he takes to be Davidson’s position:

Let Mini-English be English without a chunk of vocabulary not definable in terms of the remainder: English minus its color vocabulary, say, or minus the vocabulary of set theory… There could be speakers of Mini-English (who, we may suppose, lack the conceptual repertoire to understand English completely). This shows that there could be conceptual schemes that are subsets of our scheme. The next step of the argument attempts to show that there could be a scheme that stands to ours as ours stands to Mini-English—a superset of our scheme… Davidson’s argument to the contrary can’t be right—it could be reproduced by a Mini-English Davidson, and the conclusion would be false.76

Certainly some logically possible systems of concepts are “supersets” of ours, just as ours is a “superset” of various others. But precisely because these systems are so related, they can’t be intelligibly described as distinct conceptual schemes. Nevertheless, speakers of Mini-English may, as Byrne puts it, “lack the conceptual repertoire to understand English completely”: they may be color-blind, brain damaged, etc. Whatever the reason, however, it can’t be that their language or concepts are incommensurable with those of English speakers. Since every English speaker is also a Mini-English speaker, it can’t be that merely to make use of the one language or system of concepts would somehow make the other unintelligible.

In the same vein, Robert Kraut points out that we seem to be able to imagine “an outlaw community of Rortyan eliminative materialists, who have systematically purged their language of mental predicates” or of “Rylean behaviorists who lack the concept of an inner private mental episode”.77 He takes this to suggest the possibility of differences in conceptual scheme, as “translation would take us from their language to a proper sub-part of our own”, and Davidson has not shown that “the envisaged languages cannot endure as autonomous wholes”.78 But whether or not there could be speakers so different from ourselves that they could, somehow, consistently fail or refuse to speak of mental

---

78 Ibid., 411.
states or events, a language that is, as Kraut says, “a proper sub-part of our own” can’t be an example of an incommensurably different language. If we can imagine such speakers, then, it would seem that we are really imagining some very extreme stupidity, confusion or eccentricity, and not an alternative conceptual scheme.

Like Hacker, Byrne and Kraut describe situations in which two “languages” or systems of concepts are commensurable—by hypothesis—and then postulate some non-conceptual barrier to understanding. These commentators thus seem to impute to Davidson the absurdly radical thesis that the very idea of a failure of understanding is unintelligible, i.e., that there could not be situations in which one speaker is unable to acquire or make sense of another’s language or concepts. Similarly, Scott Soames thinks that Davidson’s argument is directed against any scenario in which we might wish to say that certain creatures “seem to have a language, but it appears to be one that we can’t interpret”:

Imagine that a group of aliens from outer space land on earth… We notice that they produce sounds that seem to allow them to communicate with one another. We further notice that they appear to gain information by looking at screens with symbols on them, and that doing what appear to be calculations with their symbols, and exchanging the results with others of their kind, seems to allow them to coordinate their activities over long distances, and to interact with natural forces in ways that are remarkably successful. In spite of all this, their symbols remain inexplicable to us… Is this sort of scenario possible or not? On the face of it, it doesn’t seem to be out of the question. Nevertheless, Davidson gives an argument designed to rule it out.79

And Neil Tennant writes that “Davidson appears to have overlooked the possibility that there might be non-human creatures with different receptivities and sensibilities” whose language is “inaccessible, in important respects, to human beings”.80

Now, if Davidson really holds that we do not even understand this kind of story, he is clearly wrong. Since there are many passages in his writings that support this

---

reading, it may be that I am disagreeing with him on this important point. Purportedly
different schemes, writes Davidson, “are variants or features we can explain to one
another, or could, given enough time, adequate attention, and sufficient intelligence on
both sides”; these are not “culture[s] where creatures communicate in ways we are
permanently disabled from penetrating”. My view is that cultures or languages
impenetrable to us are possible, but that these kinds of conceptual differences are not
sufficient for a difference in conceptual scheme. In any case, on the reading I find more
interesting, Davidson’s view is merely that this kind of story (e.g., Soames’ or Tennant’s
stories) can only be intelligibly explained in certain ways. This kind of story can’t be
explained by appeal to the idea of conceptual incommensurability, because (on the
present account, at least) this idea makes no sense. But surely this is not the only
possible explanation.

As a way of dramatizing this point, it is helpful to consider Susan Hurley’s
distinction between ability and capacity: a person is able to do X if it is possible that he
do X, and capable of doing X if it is possible that he is able to do X. Ability implies
capacity, but capacity does not imply ability. So, for instance, I have the capacity,
although not the ability, to speak Cantonese, but I lack even the capacity to remember the
names and faces of every person in China.

There seem to be many logically possible languages that we lack even the
capacity to understand. For example, there could be thinkers vastly more intelligent and
knowledgeable than any human being, whose cognitive relation to us is something like
our relation to our earliest proto-human ancestors. We are not just unable to understand
the languages of these (possible) speakers, but incapable of doing so. Or, alternatively,
there could be speakers who take centuries to pronounce a single word, or speakers with
sense organs radically unlike ours whose language contains observational terms that can
be learned only by acquaintance with phenomenal properties we can’t experience. Our
distant descendants could evolve into speakers of the Rortyan eliminative materialist
language that Kraut imagines, who come to find our writings and recording unintelligible.
Or, as Tennant suggests, there could be aliens whose communications consist of

81 “Seeing Through Language”, 128.
“emissions of variable concentrations of pheromones, along with electromagnetic signals of frequencies that our own eyes and skins cannot detect”.\textsuperscript{83} Any of these apparent logical possibilities might well provide an intelligible elaboration or explanation of the kind of scenario Soames’ describes.

In this way, at least, it would seem fairly easy to make sense of the idea of “a language we could not translate at all” (VICS, 192). We can even imagine that there are or have been actual human beings whose language is not in any sense intelligible to us. (Whether it would be warranted to hold that there are such people is, of course, an altogether different matter.) However, this is not to say that we can make sense of the idea of a difference in conceptual scheme, and many commentators seem confused on this point. The key point is that all of the intelligible possibilities envisioned just now involve non-conceptual limits on the capacities of speakers, whereas a difference in scheme would involve a special relation between concepts.

Therefore, it seems to me that the possibilities considered so far have no bearing on the intelligibility of the scheme idea, since it might be that in all such cases the languages or systems of concepts that we would lack the capacity to understand are nevertheless fully commensurable with those we do understand. Moreover, if what I take to be Davidson’s reasoning is sound, it is necessarily true, given the nature of conceptual or linguistic organization, that these systems are commensurable. This might be true even if there are many logically possible situations in which we would be unable to discover the semantic relations between languages. Conceptual inaccessibility is one thing, incommensurability another.

VI

A number of commentators have argued that conceptual inaccessibility, even without incommensurability, is enough to warrant talk of a difference in “conceptual scheme”. After all, what is really at issue here is the possibility of understanding between speakers. So perhaps my focus on the purely logical, conceptual or semantic relations between languages is mistaken. Instead, what matters is that, as I’ve allowed,

\textsuperscript{83} “Logic and Conceptual Schemes”, 77.
there could be situations in which speakers are incapable of understanding each other’s utterances. Bjorn Ramberg, for instance, argues that “incommensurability”, in Kuhn’s sense of the word, does not imply that translation or interpretation is impossible even in principle, or that truth is relative to a conceptual scheme, theory, etc., and that incommensurability, so understood, is perfectly intelligible. According to Neil Tennant, all that is needed to warrant talk of alternative schemes is the possibility of “conceptual landscapes” that would be “inhospitable” to or “uninhabitable” by human beings. In this section I will briefly consider the possibility that there might be something worth calling a “difference in conceptual scheme” without incommensurability.

Dorit Bar-On defends this kind of view as concerns the relation between English and Walbiri, an Australian aboriginal language that contains no number terms except for words we would translate by “one”, “two”, “a few” and “many”. She concedes that English and Walbiri are commensurable, in that they are “inter-interpretable”, but writes that “this does not detract from the philosophical interest” of the fact that “right now Walbiri mathematical thinking is significantly different from ours”, and that “at present our system of mathematical concepts is inaccessible to the Walbiri speakers”.

Bar-On summarizes her position as follows:
Proponents of conceptual relativism cannot rest with a claim of [conceptual] diversity. To reach a relativist conclusion, they would still need to convince us that we cannot adjudicate between alternative conceptual schemes, and, finally, that conceptual correctness is ultimately relative to a culture (or society, or theory). The radical conceptual relativist has the additional burden of arguing that the worlds in which possessors of alternative conceptual schemes live are themselves different. We have done nothing to help the conceptual relativist with all these further tasks. Our aim has been to increase the intelligibility of the conceptual diversity claim, without giving up on its philosophical excitement.

This passage is somewhat unclear as it stands. First, it is presumably not mere conceptual diversity that Bar-On takes to be philosophically interesting, but rather the

---

85 “Logic and Conceptual Schemes”, 78.
86 “Conceptual Relativism”, 164.
87 Ibid., 166.
fact that some conceptual differences are great enough to make the beliefs and concepts of one speaker inaccessible to some other speaker. Second, let us stipulate that there are differences in “conceptual scheme”, in some anodyne sense, wherever there are differences of this kind. This seems to be Bar-On’s use of the phrase.

The philosophical interest of conceptual diversity, according to Bar-On, seems to be that differences in “conceptual scheme”, so understood, might provide some kind of support for conceptual relativism—the view that we cannot “adjudicate between alternative conceptual schemes”, that “conceptual correctness is ultimately relative to [the conceptual scheme of] a culture”, that “possessors of alternative conceptual schemes” live in different “worlds”, etc. Given enough diversity, there will be inaccessibility, and inaccessibility somehow encourages these more radical claims. (This seems to be the point of her reference to “further tasks” for the conceptual relativist, for example.) The radical claims do not directly follow, perhaps, from claims of inaccessibility, but this latter claim is some kind of first step, at least, towards “a relativist conclusion”.

Now, if anything, Bar-On understates the extent of conceptual diversity among actual human beings. Consider our relation to the earliest creatures we would be willing to call human. For such creatures to understand our beliefs and experiences, they would have to undergo “retraining in another culture and/or language” so extensive that they would, arguably, at least, have to become altogether different people. These are conceptual differences even more extensive and profound than those that Bar-On discusses; although, of course, our view of the world would be radically impoverished if we couldn’t think about numbers greater than two. I think it is safe to say that we can scarcely imagine the experience or perspective of these people, and they certainly cannot imagine ours. Unlike them, we can at least identify the many ways in which their view of the world is impoverished by comparison with our own. But it is very doubtful whether this enables us to imagine what it would be like to be such a person.

I agree that if this kind of conceptual inaccessibility provides any support at all for the relativist conclusions that Bar-On mentions, there is a major lacuna in Davidson’s discussion. He does not consider the possibility of non-trivial conceptual differences without incommensurability (i.e., differences that allow for “inter-interpretability”) that

88 “Conceptual Relativism”, 164.
nevertheless provide some support for conceptual relativism. Such differences, if possible, would presumably deserve to be called differences in “conceptual scheme”, regardless of how exactly they are to be explained. (And regardless, also, of what kind of support they provide for conceptual relativism—whether the fact that there are such differences is a premise in an argument for conceptual relativism, or merely serves somehow to clarify the content of conceptual relativism.)

Bar-On assumes without argument that inaccessibility and relativism are related in this way. She assumes, that is, that some degree of conceptual difference, if it could be shown to be possible, would provide some kind of support for conceptual relativism. But it seems to me that, on the contrary, conceptual relativism is inconsistent with Bar-On’s concession that the “schemes” in question may be “inter-interpretable”. Surely there is no question that “Walbiri mathematical thinking”, for example, is not just “significantly different from ours”, but that it is also significantly inferior to ours. Merely to understand the differences is to recognize that their mathematical thinking is extremely primitive by comparison with ours. Indeed, the fact that their mathematical thoughts and language are radically inferior to ours is what explains the fact that our mathematical thinking is inaccessible to them. When we translate their mathematical discourse into our language, we realize that the whole range of their mathematical thought is identical with the most rudimentary features of our own.

This fact hardly encourages the relativist conclusion that we are not able to “adjudicate” between their mathematical “scheme” and ours. On the contrary, this kind of “significant difference” implies that their mathematical “scheme” is inferior to ours, and that we can know that it is. This is because their “scheme” is simply a very impoverished fragment of our own. Indeed, if Bar-On’s description of it is accurate, the Walbiri “conceptual scheme”, for mathematics, at least, is really just that of a child in our own culture: it is the “conceptual scheme”, more or less, of a typical Canadian 3 year old, or perhaps a severely retarded adult. This is not a polite thing to say, but when such differences are presented as evidence of some kind of special and philosophically interesting conceptual difference, it has to be said that we are already familiar with exactly these kinds of differences within our own culture or “conceptual scheme”. These
are just the kinds of differences that we normally take to be evidence merely of ignorance or low intelligence, and not of some more philosophically interesting phenomenon.

Of course, a 3 year old does not undergo “retraining in another culture and/or language” when he starts learning how to count to 10 or do simple addition. He simply learns rudimentary mathematics. Similarly, were the Walbiri to learn about counting, addition and subtraction, and so on, it would be misleading to say that they had been “retrained”. For this latter claim suggests that they had already been “trained” to think about numbers and mathematics in some way other than ours, when really they simply had not been trained to think about these things at all. At most, they had been trained to think about these things in exactly the way that we do, but only in a very limited way. There is no basis for the suggestion that, prior to this “retraining”, they already conceived of similar mathematical facts or objects in some other, conceptually alien way.

Likewise, Bar-On’s description of the relation of Walbiri to English seems inconsistent with the relativist claim that “conceptual correctness” is “relative to a culture” or “conceptual scheme” (whatever exactly that might mean). If the Walbiri were to deny that there are numbers other than one and two, for example, or if they were to claim that talk of zero or prime numbers is meaningless, these claims would not be “conceptually correct” or true “relative” to their culture or scheme. They would be simply false. Or, perhaps, we might say that they had simply failed to make any claim, because they lack the conceptual resources to have thoughts about these things.

Certainly the mere fact that they say that there is no such thing as the number 10, for example, would not support the conclusion that this obviously false claim is in some sense “correct”. Neither would the fact that their culture or scheme is “significantly different from ours”, or that our mathematical thinking is inaccessible to them. Precisely our mathematical thinking is inaccessible to them, while theirs is entirely accessible to us, we could only conclude that they are ignorant, or deeply confused. After all, we can hardly credit them with opinions about mathematics if we know that they lack the relevant concepts—much less opinions that are, in some special sense, correct or true. At least, one could maintain this view only by allowing that 3 year olds also have opinions about mathematics that, while inconsistent with those of the mathematicians, are “conceptually correct” in some special sense.
Finally, the inter-interpretability of English and Walbiri seems inconsistent with the relativist claim that speakers of Walbiri live in one “world”, and we live in another. We understand their numerical or mathematical thinking only because we can identify their beliefs and concepts by reference to a common ontology. This implies that, in this respect, the Walbiri live in our “world”, although they are ignorant of many aspects of that world. Or, if the “world” that they inhabit is the “world” that they are able to conceptualize, we should say instead that they inhabit a region of our world. These conclusions are based on our interpretation of their language, since, as Davidson emphasizes, “inter-interpretability” requires a common ontology.

In conclusion, the phenomenon of (contingent) conceptual inaccessibility without incommensurability does not suggest an argument for conceptual relativism, or even a way of clarifying the content of conceptual relativism—e.g., the notion of truth-relative-to-a-scheme, of different “worlds” associated with conceptual differences, and so on. Nor does it bear on any of the other philosophical positions that Davidson associates with the scheme/content distinction: representational theories of mind and language, Cartesian theories of an essentially private mental realm, correspondence theories of truth, empiricist theories of knowledge and meaning, etc. Absent some additional argument from non-conceptual inaccessibility to these other kinds of conclusions, there is no reason to think that Davidson has here overlooked a viable and philosophically interesting characterization of conceptual schemes.

VII

In this chapter and the last, I have defended a certain reading of Davidson’s attack on the accounts of differences in conceptual scheme based in the notion of conceptual organization. I’ve also discussed some more general and basic issues:

(i) the role, if any, of verificationism in Davidson’s discussion of conceptual schemes;
(ii) the distinction between translation and interpretation;
(iii) the question of how to individuate languages;
(iv) the distinction between conceptual inaccessibility and incommensurability.
I’ll summarize very briefly what I’ve said about Davidson’s attack on this particular formulation of the scheme idea and about these larger issues, which will figure occasionally in later chapters as well.

The idea of conceptual or linguistic *organization* is important because it might be thought to provide a characterization of language that does not imply or depend on the possibility of translation into our language. Given such a characterization of language, we could hope to clarify and defend the idea of a difference in conceptual scheme, and so to establish the intelligibility of the “very idea” of a conceptual scheme. On the reading I prefer, Davidson’s attack on the idea of conceptual organization does not depend, as many commentators seem to think, on some suppressed verificationist premise. Rather, it is directed squarely at what would seem to be the most plausible formulation of the conceptual relativist’s thesis, that we can make sense of the idea of divergent schemes (*qua* systems of organizing principles) merely by imagining that two languages stand in the same kind of relation to a single plurality of objects. This does not require that we know, in addition, which objects these are or which *predicates* (or other referential devices) each language contains.

Davidson argues that in this kind of situation, it is always possible to interpret both languages by means of a single language. This is implied by the assumption that they share a single ontology, although not, of course, a single system of concepts. The argument is not verificationist, but rather an elaboration of his earlier observation that “the dominant metaphor of conceptual relativism, that of differing points of view, seems to betray an underlying paradox” (VICS, 184). As we’ve seen, we can make sense of the idea of conceptual organization, and also the idea that different languages are like “points of view”. But this metaphor is intelligible only on the assumption that these languages are referentially related to a common ontology. But given that assumption, these languages must be translatable into a single language—a language in which the common ontology, and the relations between that ontology and each of the two languages, can be described. Putative differences in conceptual scheme thus turn out to be relatively minor, shallow differences between fragments of a single language. They are not evidence of some insuperable conceptual barrier between speakers.
In defending this view, I’ve qualified Davidson’s explicit remarks in some important ways. First, I’ve abandoned Davidson’s explicit references to *translation* in favor of the more relaxed notion of interpretation. This seemed the more natural choice of words, and consonant with Davidson’s broader theory of mind and meaning, for which the concept of interpretation, not translation, is essential. However, as I’ve said, I am not sure that there is ultimately a principled distinction to be drawn between the activities of translation and interpretation. But if there is some important distinction between the two, we should focus on interpretation rather than translation, since talk of conceptual schemes concerns the relation between language and conceptual thought.

I’ve also allowed that there may be no clear criterion of identity for languages, and so a more precise statement of the position I defend should not make any essential reference to the concept of a *single* language. The more precise formulation, then, is that on any characterization of language, translation or interpretation, if one speaker is incapable of understanding another’s words—i.e., unable even in principle to learn what the other speaker means—the reason cannot be purely linguistic. One speaker could not be incapable of understanding what another speaker is saying merely because their languages (however individuated) are different. Of course, this thesis is considerably weaker than Davidson’s contention that “translatability into a familiar tongue” is “a criterion of languagehood” (VICS, 186). But then, so is his remark, in a slightly later paper, that his view “does not make the ability to interpret a language depend on being able to translate that language into a familiar tongue”. 89

Finally, I’ve argued that while failures of understanding cannot be explained in a way that would justify or clarify talk of “conceptual schemes”, there could nevertheless be “a language we could not translate at all” (VICS, 192). Our capacities are finite, and it is at least logically possible that there are speakers whose communication is beyond our capacity to recognize or understand in various ways. These speakers would be conceptually *inaccessible* to us, even though their languages or concepts would be commensurable with ours (on the present account of incommensurability, at least). This

89 Donald Davidson, “Reply to Foster”, in *Inquiries into Truth and Interpretation*, 175: Davidson points out that one can understand or interpret another’s utterance of “I’m warm”, for example, without being able to translate that utterance into one’s own language or idiolect. The interpreter’s (correct) statement of truth conditions will not have the same meaning as the utterance he is interpreting, but it is nevertheless sufficient for interpretation.
much is intelligible. However, this concession does not suggest any new defense or clarification of conceptual relativism.

If, as some commentators hold, incommensurability does not require that interpretation is impossible across languages for purely conceptual reasons, but is merely a kind of contingent conceptual inaccessibility, the notion of incommensurability is intelligible. But on this view, it isn’t clear why we need this notion, since it is agreed on all sides that interpretation can sometimes be very difficult or awkward; interpretation may even, as a matter of contingent fact, prove impossible in some cases. We might discover that the Walbiri are simply unable, as far as we can tell, to learn mathematics. There would no point in saying that they have a different “conceptual scheme”, unless that is just a polite way of saying that they are less intelligent than we are. The uncontroversial fact that interpretation can be difficult or awkward is of no philosophical interest. (The conceptual differences between individuals or groups that sometimes make interpretation difficult or awkward may, of course, be quite interesting in other ways.) Nor is it of any philosophical interest that some thinkers may be unable, for various familiar non-conceptual reasons, to acquire certain concepts, theories or languages.

It might be objected that these qualifications leave it unclear what exactly Davidson means to say, on the reading I defend. But of course this is to be expected if, as he contends, the very idea of a difference in conceptual scheme is unintelligible. Davidson’s conclusion is that we can’t “give a solid meaning to the idea of conceptual relativism” (VICS, 197), and that talk of conceptual schemes is thus equally nebulous. If this is right, there can be no precise statement of what it is that, according to Davidson, is unintelligible. We can only examine the claims and concepts associated with the scheme idea and discover that, on reflection, there does not seem to be anything worth saying about so-called “conceptual schemes” that would not be better expressed in some other way. Of course, we have not yet established this conclusion. And the lack of a precise formulation of the thesis I’ve ascribed to Davidson is not a virtue of his position (or my discussion of what I take to be his position). My point is merely that if Davidson is right to deny the intelligibility of the scheme idea, this lack of precision is inevitable.
CHAPTER 4
ORGANIZATION, PART 2: INTENSIONAL DIFFERENCES

The last chapter focused on an extensional construal of the metaphor of organization. Languages “organize” a single plurality of objects differently, on this reading of the metaphor, if the extensions of their terms are different. However, it might be objected that this overlooks the possibility that intensional differences between languages might prevent translation or interpretation. For one might hold that two languages “organize” a single plurality of objects differently if the meanings or senses of terms in the one language are different from those of terms in the other. This would surely constitute a difference in ways of organizing, classifying or referring to objects, regardless of whether the terms in question are co-extensive. Moreover, this is a natural reading of the claim that it is concepts that differ across conceptual schemes. Presumably the concept of a kingdom, for example, is precisely the meaning or sense of the word “kingdom”—at least, these entities would seem to be quite closely related.

Davidson would reject any appeal to differences in intension as a basis for talk of conceptual schemes for the radical reason that, in his view, senses, propositions and other intensional entities play no role in a theory of linguistic meaning:

Paradoxically, the one thing meanings do not seem to do is oil the wheels of a theory of meaning … My objection to meanings in the theory of meaning is not that they are abstract or that their identity conditions are obscure, but that they have no demonstrated use.90

According to Davidson, theories of meaning for natural languages are purely truth conditional. Nothing more is needed to “give the meaning”,91 so to speak, of expressions in a language. Since meanings, senses, propositions and the like are introduced into our ontology only as theoretical posits of (bad) theories of meaning, we have no reason to believe in such things.

91 Ibid., 20.
Thus, when Davidson claims that there could not be a language that is untranslatable into our language he means to deny that there could be a language for which it is impossible in principle to devise a theory of truth conditions expressible in our language. (Or, more precisely, on the reading I defend, he holds that there could be some language or other by means of which the truth conditions of both languages could be stated. That language need not be ours.) He does not mean that the sense or proposition expressed by a sentence in any language must be expressible by means of any other, since he sees no point in positing such entities to begin with. Obviously, if intensional entities explain nothing about the meaning of any language, intensional differences between languages will not explain how the meaning of one language might be uninterpretable by means of another.

The objection from the possibility of intensional differences, then, is not one that Davidson would deem worthy of serious consideration. But it would be fair to question Davidson’s assumption that his austere and controversial notion of meaning, and hence translation, is a reliable measure of identity or difference of conceptual scheme. In what follows, I will assume for the sake of argument that we have some understanding of the intensional notions that Davidson rejects, and that to really or fully understand a language requires not only knowledge of a (true) truth theory for that language, but also knowledge of the sense or concepts associated with expressions of the language. So it is not enough to know that an alien sentence $p$ is true if and only if a sentence $q$ in one’s own language is true. In addition, one must know that $p$ and $q$ express the same propositions (and that their component terms express the same concepts). If inter-translatability is the criterion for identity of conceptual scheme, we should then say that two languages belong to the same scheme if it is possible to provide a sense preserving translation of either by means of the other.

In passing, we should note that the ability to produce a sense preserving translation does not seem to be a necessary condition for interpretation or understanding. Consider a 1980 token of the sentence “Now it’s raining”, for example. It seems that we fully understand the sense of that token, even though there is nothing that we can (now) say that has precisely that sense. The same is true of all sentences containing indexicals. Although we can state their truth conditions, it will always be possible for someone to
hold true the object language sentence but not the meta language sentence stating its truth condition, and vice versa. Thus, if understanding a sentence consists in a kind of acquaintance with the proposition it expresses, it seems that we can be acquainted with propositions we can’t express. Or, if acquaintance with a proposition requires the ability to express that proposition, such examples seem to show that we can sometimes understand a sentence without being acquainted with the proposition it expresses.

But let us suppose for the sake of argument that, with the possible exception of indexical language or thought, the possibility of sense preserving translation is necessary for understanding, and hence for identity of conceptual scheme. As before, “translation” is to be understood broadly, such that the possibility of a sense preserving translation of two languages by means of a third would make the first two inter-translatable. Schemes differ, then, to the extent that such translation is impossible. If such translation is totally impossible across certain languages, those languages are associated with totally different conceptual schemes. In the next section, I’ll try to clarify this proposal. Then I’ll argue that the conditions necessary for a difference in conceptual scheme, so understood, could never be satisfied by any pair of languages or systems of concepts.

I

If the possibility of sense preserving translation is necessary for identity of conceptual scheme across languages, what kind of intensional difference might be sufficient to prevent such translation? As we’ve seen, mere conceptual difference, no matter how radical, is not sufficient for a difference in conceptual scheme. If an alien term expresses a concept C true of cabbages, but utterly unlike any concept of ours and unrelated to any of ours, it would be possible, in principle, to simply add C to our own system of concepts.

Of course, it might be that, for some non-conceptual reason, we aren’t able to understand the more comprehensive system that comprises both our concepts and the alien concepts (i.e., to learn or invent a language that affords a sense preserving translation of our language and the alien language). But in such cases, as we’ve seen, the inability of speakers to understand an alien language is not sufficient for a difference in
conceptual scheme. If schemes differ, the impossibility of conjoining two systems of concepts must be due simply to some special relation—“incommensurability”—that holds between those systems.

What we are trying to imagine, instead, is that merely to conceive of things in the way that we do makes it impossible to conceive of those same things in some alien way. If two systems of concepts are related in this way they must be in some kind of tension or conflict, and not merely different. Otherwise, the two systems can, in principle, be conjoined.

And yet, although the two systems must be in some kind of conflict, the conflict must be something other than cognitive disagreement. We are in a familiar kind of conceptual conflict if we conceive of the same things in incompatible ways—e.g., if I think of John as a bachelor and you think of him as a husband. But this kind of conflict is simply disagreement concerning a proposition entertained by two thinkers: you affirm the proposition that John is married, and I deny that same proposition. Alternately, we might say that disagreement concerns the extension of a shared concept: you hold that John is in the extension of the concept *husband*, say, and I hold that he isn’t. This familiar kind of disagreement is possible only if there is mutual understanding, which is precisely what a difference in conceptual scheme is supposed to preclude.

So, to summarize, there would seem to be (at least) two necessary conditions for a difference in conceptual scheme of the kind under discussion:

1. The system of concepts associated with a language L1 is applicable to the same objects as the (different) system of concepts associated with another language L2.
2. Speakers of L1 are in some kind of conceptual conflict with speakers of L2, other than cognitive disagreement.

Condition (1) captures the idea that divergent schemes “organize” the same plurality of objects differently. The predicates, quantifiers and singular terms of L1 and L2 serve to classify, individuate or refer to a single plurality of objects by means of different concepts. This is not sufficient for a difference in scheme, since concepts associated with a single language may be related in this way. A language related to ours in this familiar way would be associated with a system of concepts that was, again, *merely* different and additional to ours. Condition (2) is thus necessary if we are to make sense of the further
claim that one language or system of concepts is incommensurable with another. Conditions (1) and (2) may not be jointly sufficient for a difference in conceptual scheme, since there might be kinds of conceptual conflict other than cognitive disagreement that nevertheless allow for mutual understanding. But it would seem that (1) and (2) are necessary conditions, at least, for a difference in scheme.

Now there is, arguably, a kind of conceptual conflict other than cognitive disagreement, i.e., other than the kind of disagreement that can be cast in terms of shared concepts or propositions entertained by both parties. Let us suppose that there are truths or principles constitutive of concepts, so that to acquire or deploy concepts requires an acceptance of certain truths or principles. On this assumption, it would seem to be possible for thinkers to disagree over those constitutive principles. Thus, if the truth that bachelors are unmarried is constitutive of the concept bachelor, a thinker who has the (de dicto) belief that some bachelors are married would be in conflict with thinkers who accept the principles constitutive of the concept.

It might be objected to this suggestion that if some principle P really is constitutive of a concept C, it is impossible for anyone to make use of C while denying P—to say that this is possible is just to deny that P is constitutive of C. This objection is unconvincing. According to this objection, the fact that there are meaning- or concept-constitutive principles implies the psychological impossibility of confusion or irrationality. But this doesn’t follow. In order to make sense of this special kind of conceptual conflict, we need only the plausible assumption that we are not logically and conceptually omniscient, that we are sometimes ignorant or mistaken about some of the inferential relations between principles constitutive of a concept, or between those and other principles constitutive of other concepts, without thereby losing the relevant concepts altogether. Or, at the very least, let us assume that we could be in that position at times, which is all that the present suggestion requires.

To be sure, if all the principles \{P_1, P_2, \ldots P_n\} are constitutive of a concept C, one could not fully grasp C while denying any of those principles, but this doesn’t rule out the possibility of a partial grasp of C, sufficient for confusion. And it is precisely because confusion is incompatible with a full grasp of the relevant concept or concepts that the kind of conceptual conflict under discussion does not reduce to cognitive
disagreement. In a disagreement over constitutive principles, the two parties grasp the disputed concept in different ways or to different degrees. There is a sense, then, given the vagueness of the term “concept”, in which the concepts of the two parties might be said to differ, and a sense in which they are the same. But this is exactly what would seem to be needed for a difference in scheme. In any case, there appears to be a kind of conceptual conflict here that is not present in cases where both parties accept the same principles as constitutive of a concept, but disagree about its extension.

Obviously, however, this kind of conceptual conflict does not provide a good model for differences in conceptual scheme. A radically confused thinker would be difficult to understand, but his concepts are not in conflict with ours. These are either the same as ours or, again, they form an additional and complementary system of concepts. By his own standards his thoughts are incoherent, although he may not realize that fact. Thus, if he were less confused, there would be no conceptual obstacle to mutual understanding.

What we should try to imagine, instead, is a situation in which one thinker’s use of concepts is (largely) coherent, by his standards, but incoherent by the standards of some other thinker. Only in that case would there be a genuinely conceptual conflict between thinkers, which is not reducible to cognitive disagreement and which might make mutual understanding impossible. This seems a promising suggestion, moreover, in that it would seem to encourage conceptual relativism, i.e., the notion that what is true or real is relative to some set standards or concepts. Let us explore this suggestion in more detail.

First, notice that this situation would require that at least some concepts be shared across conceptual schemes. Otherwise, there will be no conflict and the purported difference in conceptual scheme will collapse into mere differences between complementary concepts. The suggestion is that the standards of two systems of concepts yield conflicting verdicts concerning the coherence of concepts or propositions, but these standards will not be applicable across systems unless there is some basic set of concepts common to both. So, on the present suggestion, a total difference in scheme is impossible, since that eliminate the possibility of conflict. However, this is a fairly minor concession.
Suppose that *some* concepts of an alien scheme are common to ours—concepts of elementary logic, perhaps, or concepts of simple shapes and colors. But suppose that we could not come to understand any of the others, which are not shared. The aliens’ psychological or ethical beliefs, for example, would be unintelligible to us, in principle. Indeed, it would seem to follow from these assumptions that, although the aliens have certain beliefs that are unintelligible to us, these beliefs can’t be accurately categorized in any of the ways we categorize ours—as “ethical”, “psychological”, “religious”, “scientific”, etc. This would be a very radical degree of conceptual otherness—so radical that it is scarcely imaginable. Nothing much has been lost, then, in abandoning the claim of *total* incommensurability.

In what follows, I will treat *incoherence* as a form of inconsistency. When a proposition is incoherent, it involves the ascription to an object of properties that could not possibly belong to a single object at one time. (Or, cast in terms of concepts, let us say a proposition is incoherent if it involves the application to an object of two or more concepts that could not possibly be co-extensive.) We’ll see later that this notion of incoherence is not strictly necessary for the argument I want to develop. However, this notion of conceptual conflict, and of incoherence, seems natural enough, and it affords a fairly precise characterization of an *instance*, at least, of the general kind of view under discussion.

II

Schematically, the relation between concepts that we are trying to imagine is this:

(i) Principles P1 and P2 are constitutive of some concept C belonging to system 1.
(ii) P1 is consistent with P2 by the standards of system 1, but inconsistent with P2 by the standards of system 2.

For instance, P1 might state that an object is in the extension of C if and only if it is a bachelor, and P2 might state that an object is in the extension of C if and only if it is married. Both system 1 and system 2 contain the concepts *bachelor* and *married thing*, but these shared concepts figure in the constitutive principles of concepts unique to system 1 in ways that are, by the standards of system 2, impossible. This situation would
seem to make understanding impossible, although not in the way we might have expected.

Let us briefly retrace the steps that led to the proposal now under discussion. Initially, we considered the suggestion that Davidson’s austere notion of translation should be replaced by a richer criterion for identity of conceptual scheme. On this suggestion, schemes differ where a sense preserving translation from one language into another is impossible. But if this means that it is impossible to re-express in one language *anything* that can be said in another, i.e., any of the propositions expressible in that second language, the intensions associated with the second language turn out to be merely different and additional to those associated with the first. Speakers of either language could, in principle, simply learn the other, and the putative difference in schemes collapses into mere conceptual diversity. (In that case, we could say that the two languages are “inter-translatable” in the broad and no doubt misleading sense in which “translation” is synonymous with “interpretation” or “understanding”. But it is crucial to distinguish *this* sense of “translation” from the more usual, narrow sense.)

So, as we’ve seen, if differences in conceptual scheme can be explained by intensional differences, some intensions must be shared across languages—i.e., some concepts, meanings, senses or propositions. Hence a partial sense preserving translation must be possible across schemes. The present conclusion is that a *full* sense preserving translation (in the narrow sense) must be possible across schemes, but that most or some of what can be expressed in both languages is *intelligible* only to speakers of one of the two. For example, we are able to express in our language the proposition that John is a married bachelor, and there is some sense in which we can understand that proposition. But precisely because we can understand it in a certain sense, we find it incoherent, and hence unintelligible, in some deeper sense. This proposition involves the application to an object of concepts that not only do not but, by our standards, *could not possibly apply* to the same object (at one time).

Surprisingly, then, the criterion of inter-translatability turns out to be superfluous as concerns the present account of conceptual schemes. For on this account, a full sense preserving translation will be possible for languages associated with divergent schemes just as it will be for languages associated with a single scheme. Instead, the conclusion at
present is that schemes differ where speakers of one language are able to express in that language whatever can be said in some other, but are unable, in principle, to understand what they can express. Translation has turned out to be no measure of understanding or interpretation.

The idea of an unintelligible sense, or of a sense preserving but unintelligible translation is not as strange or paradoxical as it may seem. The proposition that $1 + 1 = 3$ is surely intelligible, in some sense, but also incoherent, and so unintelligible in some other sense. Only because we understand that proposition in some sense can we realize that it is incoherent, and therefore unintelligible in a deeper sense. In any case, let us simply grant for the sake of argument that there is some distinction, whatever exactly its nature, between kinds or degrees of understanding, which is illustrated by this example. This is the distinction involved in the present account of differences in conceptual scheme.

Even granting all of this, however, it is hard to see how both of conditions (i) and (ii) could be satisfied by any pair of languages or systems of concepts. For what it could it mean to say that certain propositions consistent by (or relative to) the standards of one system and inconsistent by (or relative to) the standards of the other? Of course, this cannot mean merely that those who make use of the first system believe that certain propositions are inconsistent (by their standards). That might be true even if those propositions are consistent, in which case the putatively incommensurability between systems collapses into a mere conceptual difference. That is, if C is the concept of a married bachelor, and if it is consistent to say that something is (at once) both married and a bachelor, then C is a concept that we can simply add to our system of concepts. Once again, it turns out that there is no conflict between C and any concepts we have, although we might mistakenly believe that there is.

We may be psychologically incapable of understanding how the constitutive principles of C could both apply to a single thing at one time. But this is again merely to suppose that some non-conceptual factor makes it impossible for us to give an interpretation or sense preserving translation of certain sentences. What is necessary for a difference in scheme is instead some condition to the effect that our concepts make such translation impossible, not only for us, but for any thinker making use of our
concepts. This reading of condition (ii) thus fails to capture this key feature of a difference in conceptual scheme.

So presumably condition (ii) must mean that the disputed propositions are *in fact* consistent by the one set of standards, and yet inconsistent by the other set of standards. But this, in turn, implies that there is some system of logical standards or principles more general and basic than those of either system, in terms of which both can be evaluated. For the idea is not merely that certain propositions have logically unrelated and hence compatible properties—e.g., that the constitutive principles of some concept C are, so to speak, “system-1-consistent” and “system-2-inconsistent”. If so, the purported difference in scheme collapses again into a mere difference between complementary concepts. Rather, both systems must share a single concept of consistency that can be applied to either system, with conflicting results. Since this can’t be the concept of system-1-consistency or system-2-consistency, it must be some prior concept of consistency.

Therefore, on the present account of conceptual schemes, there must be a system of logical principles or standards more general and basic than the standards of system 1 or system 2, which can be used to evaluate the logical relations between the standards of both systems and the disputed propositions. When cast in terms of system 3, the standards of the first two yield inconsistent verdicts on the consistency of P1 and P2 (and other propositions). One set of standards implies that P1 and P2 are inconsistent, while the other does not. But this is logically impossible if the terms “imply” and “consistent” are being used univocally.

The claim that some proposition *p* is consistent by or relative to the standards of system 1 can only mean that *p* could be true *if* certain (normative) propositions of system 1 were true. Obviously, however, if *p* could not even be *true*, by the logically prior standards of system 3, it could not be true *together with* some other set of propositions—i.e., those stating the logical standards of system 1. So *p* must be consistent in this prior sense if it is consistent *relative to* either system. But if it is consistent in this logically prior sense, no proposition or set of propositions of system 2, which is logically subordinate to this prior set of standards, can imply that it is *not* consistent. Of course, certain propositions of system 2 might imply that *p* is false, but in that case the difference between schemes collapses into a cognitive disagreement.
To make an analogy, suppose that Banff and Jasper pass different laws concerning loitering. Now, either there is some provincial, federal or constitutional law against loitering, or there is no such law. If there is some such law, it is incoherent to say that loitering is legal in Banff but illegal in Jasper: it is illegal in both places, and hence the Banff “law” is itself illegal in this more general and basic sense. On the other hand, if there is no such law, both municipal governments have the authority to deal with loitering as they choose. In that case, it may be that loitering is legal in Banff but illegal in Jasper. Then there is no conflict between the laws of the two cities, since the laws of Jasper don’t apply to what goes on in Banff. In other words, it is really loitering-in-Jasper that is illegal; or, perhaps, we could say loitering is illegal-in-Jasper. Thus, if there is to be a legal conflict, there must be some form of legality that applies to the laws of both cities, such as provincial law, say. But if there is, it will also turn out that at least one of those municipal laws is itself illegal in this latter sense, and so, in that sense, not a law.

III

The present account of conceptual schemes thus seems to be incoherent, because it involves inconsistent assumptions. On the one hand, there must be certain logical principles governing divergent schemes, or else there will be no single notion of consistency deployed in conflicting ways across schemes, and the purported difference in scheme will collapse into a mere difference between complementary concepts. On the other hand, there cannot be any such principles, for if there are, any propositions consistent by the standards of either scheme will be consistent by those of the other scheme also, and the two will again collapse into one.

The general line of argument developed here does not require any assumptions as to which logical principles are genuinely basic and universal, or even the assumption that there are such principles. If there are no ultimate logical principles applicable to all theories, languages or systems of concepts, then putatively different schemes are complementary systems of concepts, each system having its own logic and its own concept of consistency. Whatever it might mean to call all of these “logics”, or concepts of consistency, there is no conflict between any systems so related. On the other hand, if
there are ultimate and universal logical principles, then, whatever these principles might be, putative differences in scheme involve real conflicts, but these are just cognitive disagreements.

Notice also that the argument does not even require the assumption that the kind of conflict under discussion is strictly logical—or that is “logical” in our sense of the term. For we can instead stipulate merely that certain propositions $s$ and $t$ are constitutive of a concept $C$ belonging to one system or theory, and yet “incompatible” somehow by the standards of some other system or theory. Suppose that the constitutive principles of some alien concept $C$ are that (a) an object is in the extension of $C$ if and only if it is red, and (b) an object is in the extension of $C$ if and only if it is blue. These principles are incompatible, but seemingly in some non-logical sense of the term.

Even so, the incompatibility, whatever its nature, is a real incompatibility only if there is a concept of incompatibility that is applicable to both systems, and in particular to claims about incompatibility relative to the standards of either. Again, the idea cannot be merely that we believe that principles (a) and (b) are incompatible given our color concepts, since that leaves open the possibility that the alien concepts are simply different and additional to their existing system. Nor can the relevant kind of incompatibility be relativized to the conflicting systems, for then there is no conflict between the properties of being compatible-relative-to-$A$ and incompatible-relative-to-$B$. (Or, more precisely, if there is a conflict between those relativized properties, it can’t be explained by appeal to other relativized properties—e.g., the property of being red-relative-to-our-scheme. Rather, it must be explained by appeal to some more general and basic concept of incompatibility that is not relativized to either of the purported schemes.)

Therefore, there is some prior notion of compatibility that governs both systems and, in particular, the notion of compatibility relative to either system. But then, by the same reasoning as before, the principles constitutive of concepts in either system must be compatible relative to the standards of both systems, or compatible relative to neither. Contrary to condition (ii), these principles cannot be compatible relative to just one of the two sets of standards. Hence, as before, it would seem that there must be, and yet can’t be, some more general and basic system of principles applicable to both putatively different conceptual schemes.
Finally, it is worth pointing out that proponents of the scheme idea cannot defend against this argument by denying the assumption that certain truths or principles are constitutive of concepts. For suppose that, as Quine holds, there are no such truths or principles, i.e., truths or principles that, unlike ordinary empirical propositions, merely state the contents of concepts or the meanings of words. In that case, every putative difference in scheme can be seen as a cognitive disagreement. Some people may wrongly believe (de dicto) that certain things are married bachelors, but the reason they are wrong can only be that, in fact, nothing is both married and a bachelor. However, the reason can’t be that the concepts bachelor and married thing are conceptually incompatible, i.e., that nothing could possibly be (simultaneously) in the extensions of both. That would be possible only if there are principles constitutive of those concepts that somehow exclude one another. On the Quinean view, there are no purely conceptual facts, no facts about which concepts we have that can be separated from the facts about which empirical beliefs we have. If this is right, then can be no purely conceptual differences between thinkers. So there could not be, in particular, a special kind of purely conceptual difference that makes interpretation impossible.

V

The argument developed in this chapter depends on the presupposition that propositions must always be either consistent or inconsistent. This is what ensures the conclusion that purported differences in conceptual scheme collapse into mere diversity of consistent propositions, or complementary concepts, or else mere disagreement, i.e., cases in which one thinker believes something inconsistent with what another believes. Some philosophers may hold that this begs the question, since, as we’ve seen, the very idea of a difference in scheme appears to require that there is neither agreement nor disagreement across divergent schemes. This is just to say that the propositions associated with one scheme are neither consistent nor inconsistent with the propositions associated with another.

Thus, Carol Rovane describes the kind of argument I have advanced against the intelligibility of the idea of a difference in conceptual scheme, which she thinks is not
Davidson’s, as being based in a “formal” difficulty concerning characterizations of the “exclusionary relation” that is supposed to hold between divergent schemes:

Many philosophers hold that there is no way in which claims can exclude one another except by being logically inconsistent. They insist that different schemes must have a common subject matter over which they disagree because, otherwise, nothing would stand in the way of embracing them together in much the same way that we can embrace theories about disparate phenomena such as geography, botany, and musicology.

She proposes that this objection can be met by thinking of the special “exclusionary relation” between schemes as

a kind of restriction on truth-functional relations, so that truth-functional relations can hold only within schemes but not across them. This would imply that claims belonging to alternative schemes would be neither consistent nor inconsistent.

The upshot is that the more basic disagreement between those who claim to understand the scheme idea and those who claim it to be unintelligible concerns the question of “whether truth functional relations hold among all truth value bearers”.  

Rovane points out that the assumption merely that there is a single, complete set of truth makers does not entail that these “are bound by the normative constraints that are implicit in the absolutist ideal of a single and complete body of truths, most notably, consistency”; after all, “a set can contain anything”. Therefore, it begs the question to presuppose that any two true propositions must be consistent. Some further reason is needed, then, for the “absolutist” claim that if certain propositions that are supposed to be associated with divergent schemes are true, they must be consistent with one another and, therefore, not really associated with divergent schemes at all.

As this suggestion applies to the present formulation of the scheme idea, we should then consider the possibility that the principles constitutive of one set of concepts are, in this way, isolated from those associated with some other set of concepts. The two sets of principles are neither consistent nor inconsistent. It is interesting to notice that, on this suggestion, we can also drop the requirement that there be some base set of concept

---

93 Ibid., 115.
common to both schemes, in relation to which some kind of conflict could arise. For perhaps the most basic principles of one scheme are formulated by means of basic concepts that are simply different from those that figure in the principles of some other schemes, and yet the two sets of basic principles are not consistent (or inconsistent).

I think this suggestion is implausible, however, on any account of truth. The problem is that our understanding of truth cannot be intelligibly separated from our understanding of truth functional relations between vehicles of truth. What could it mean to say that two sentences, \( s \) and \( t \), are both true, while denying that they could both be true? To be sure, if we accept Rovane’s proposal, this would not be to say that they could not both be true. But surely it is just as unintelligible to say that some situation \( S \) is actual while denying that \( S \) is even possible.

Moreover, it seems to me that Rovane’s claim that “a set can contain anything” is simply false, if it is supposed to mean that any set of a particular kind could be expanded to include entities of any kind. Perhaps it is true that, for any collection of entities, there is a set containing just those entities. But a set of numbers, say, can’t be expanded so as to contain peaches or colors without ceasing to be a set of numbers. Likewise, it does not seem possible that a set of true propositions could contain propositions that are neither consistent nor inconsistent with one another, while nevertheless being a set of true propositions—or even a merely set of propositions, true or false.

It may be that Rovane’s puzzling position here stems from a mistaken view of what motivates the “absolutist” view that all propositions are either consistent or inconsistent. She believes that this view is supported by an inference from “the claim that there are mind-independent truth-makers”, which she associates with a certain response to the “Kantian problematic”. This seems confused. Let us suppose that all of reality is somehow “mind-dependent”—e.g., that when the last person leaves the room, the room pops out of existence. This (absurd) view would not give any support to the claim that some truths are not consistent with others, nor would it suggest any explanation of how that could be. The “absolutist” view is not based on the claim that what makes sentences or other vehicles of truth true or false is typically independent of any mind, but merely on an understanding of the concept of truth.

---

94 “Ethical Scheme”, 115.
In conclusion, Rovane’s proposal does not require an additional argument on behalf of “absolutism”, because it purports to explain the mysterious “exclusionary relation” between conceptual schemes on the basis of a conception of truth that is at least as mysterious as what it is supposed to explain. Perhaps it is true that there is no way to argue against the view that truth could be somehow “relative” to a conceptual scheme, and hence that certain truths are not consistent, without begging the question. But this is inevitable if this kind of view is simply unintelligible.

By analogy, there is no way to argue against the claim that nothing is true without presupposing that something is true. What one aims to do, instead, is to reveal that one’s opponent is also presupposing that something is true, or else that he has simply failed to say anything that is intelligible, even by his own standards. Similarly, we can legitimately argue that the notion of truth relative to a scheme is either incoherent, because it can be explained only on some hidden assumption that truth is not relative to a scheme, or else unintelligible. Indeed, if this kind of view is in fact unintelligible, to “beg the question” against such a view would be simply to make sense, to base one’s arguments on premises that are not themselves unintelligible.

V

This general form of argument developed in this chapter seems devastating for any account of conceptual schemes that appeals to intensional differences between languages. A workable account of that kind would require a notion of conceptual conflict that does not involve any implicit appeal to principles, standards or norms, since, if it does, it will be vulnerable to some version of the argument developed here. But it is hard to imagine how anything could count as a conflict between concepts (or systems of concepts) without being open to that kind of description. But if concepts or modes of presentation are understood in some very different way, it seems to follow immediately that these are at most very different from one another.

For instance, perhaps the concept red is a purely recognitional or phenomenal concept, its content is just a certain kind of experience that can’t be characterized by any set of application conditions or constitutive principles. And perhaps we can imagine
blind Martians who navigate by echolocation, like bats, and have some concept C that is, for them, purely recognitional or phenomenal. The content of C would then be unimaginable for us, and the content of the concept red would be unimaginable for them. But in what sense could we say that their concept is in conflict with ours?

It is also interesting to note that the argument developed here, although not to be found in Davidson’s discussion of conceptual schemes, seems to be an instance of his general strategy. As we noted earlier, Davidson’s case against the scheme idea is an elaboration of his initial claim that the idea of a difference in scheme seems to involve “an underlying paradox”:

Different points of view make sense, but only if there is a common co-ordinate system on which to plot them; yet the existence of a common system belies the claim of dramatic incomparability. (VICS, 184)

Where the metaphor of organization is taken to involve extensional differences, the common co-ordinate system is simply the plurality of objects individuated, classified or referred to by the concepts of the two purportedly divergent schemes. Here the “common co-ordinate system” is the system of shared concepts and logical or conceptual standards needed to make sense of the idea that there is a conflict between one system of concepts and another. Only if there are shared concepts can there be some kind of dispute over the coherence of other concepts, e.g., the concept of a married bachelor.

So there is a sense in which, on this kind of view, the organization of a common plurality of objects is dependent on a prior organization of these common concepts. The common concepts are themselves “organized” in that they figure in the principles constitutive of other concepts unique to one scheme or the other. However, as we’ve seen, these shared concepts and standards imply that we can compare purportedly divergent schemes, so understood, and find them to be identical. Thus, as in the extensional case, it appears that the “underlying paradox” of conceptual relativism is a contradiction.

I should perhaps forestall a radical objection to the argument developed in this chapter. According to this objection, it begs the question against conceptual relativism to presuppose the legitimacy of any of our concepts or norms as applied to an alien conceptual scheme. The very idea of a totally alien scheme, after all, is the idea of a
system of concepts that we would find it impossible to understand. Therefore, it is no argument to show that, from our perspective, any purportedly alien scheme will be simply incoherent or logically impossible, since that is exactly what we would find if the that system of concepts were truly an alien conceptual scheme.

This kind of objection is confused, since my argument does not presuppose that there could not be a system of concepts that appears, from our perspective, to be simply incoherent or logically impossible. Rather, it presupposes merely that if the idea of a difference in conceptual scheme is intelligible to us, the relation that is supposed to hold between our “scheme” and some other must be intelligible. This is not to say that the other scheme must itself be intelligible to us. Now obviously any such description must be cast in terms of “our” concepts, “our” logical or other standards, and so on. And since it must be cast in those terms, it is legitimate to argue against the coherence of such a description by appeal to those same terms.

To be sure, this strategy does not apply to the most radical formulation of the scheme idea: the view that, although alternative schemes are possible, there is no way for us to give any account of their relation to our concepts, or for us to have any reason to believe that they are possible. But if none of our concepts are applicable to a description of the relation of an alien scheme to our own, the alien scheme can’t be correctly described as a system of concepts, nor even as something that is somehow different from our system. Nor can we even say, presumably, that there is something or other that we can’t imagine or understand. The proposition that alternative schemes are possible is then not one that we could ever entertain; obviously it is impossible that some proposition is both a proposition that we can’t understand and one that some of us, certain relativist philosophers, believe to be true. So the most radical formulation of the scheme idea is self-defeating, regardless of the merits of any arguments developed here.
In the last two chapters, we’ve seen that neither an extensional nor an intensional construal of the metaphor of conceptual organization provides a basis for the idea of a difference in conceptual scheme. I now turn to Davidson’s attack on formulations of the scheme/content distinction that appeal to a second class of “images and metaphors” (VICS, 191), according to which conceptual schemes serve to predict, fit, face or account for experience. Davidson writes that this proposal seems to concern the role of complete sentences, rather than the sub-sentential “referential apparatus of language”:

It is sentences that predict (or are used to predict), sentences that cope or deal with things, that fit our sensory promptings, that can be compared or confronted with the evidence… The general position is that sensory experience provides all the evidence for the acceptance of sentences (where sentences may include whole theories). A sentence or theory fits our sensory promptings, successfully faces the tribunal of experience, predicts future experience, or copes with the pattern of our surface irritations, provided it is borne out by the evidence. (VICS, 193)

Davidson explains that, on this view, the sensory evidence “is not just actually available evidence: it is the totality of possible sensory evidence past, present and future” (VICS, 193). According to this formulation of the scheme/content distinction, then, conceptual schemes are to be identified or closely associated with groups of sentences (i.e., theories) “borne out” by all possible sensory evidence.

In this chapter, I’ll consider some features of Davidson’s discussion of this “general position”—i.e., knowledge empiricism—as a basis for an account of differences in conceptual scheme, and conclude that Davidson’s argument rests on a tendentious characterization of this position. But although I’ll raise doubts about Davidson’s reasoning, I agree that the views he associates, rightly or wrongly, with the metaphor of “fit” do not provide an intelligible account of differences in conceptual scheme. I’ll also argue that Davidson’s critique of empiricism provides a different and more powerful case
against the proposed account of differences in conceptual scheme, which does not rest on a tendentious characterization of knowledge empiricism.

I

Characterizations of conceptual schemes by reference to the metaphor of fitting experience fail, according to Davidson, because “the notion of fitting the totality of experience [or evidence], like the notion of fitting the facts, or of being true to the facts, adds nothing intelligible to the simple concept of being true” (VICS, 193-4). If some theory is “borne out” (warranted, confirmed or supported) by “the totality of experience”, or “the totality of possible sensory evidence”, that theory is true. So the notion of the totality of experience or evidence does not provide any contrast between scheme and content except whatever contrast, if any, can be drawn from the “simple concept” of truth.

But it may seem that the concept of truth does provide just such a contrast. For if a theory (i.e., a group of sentences) is true, there must be something that makes it true. Thus, even if the totality of evidence or experience will not serve as something “neutral and common” (VICS, 190) in relation to which conceptual schemes (qua theories) might be thought to differ, perhaps that neutral and common entity can be identified with whatever it is that makes sentences and theories true. Davidson argues that there is no way to make sense of this latter suggestion. To say that a sentence “corresponds” to some entity that “makes” it true is just to say (if anything) that the sentence is true.

Davidson concludes that the present account of conceptual schemes reduces to the “simple thought” (VICS, 194) that an alternative conceptual scheme is (or is associated with) a true theory that cannot be translated into our language. (Or at least, a largely true theory; Davidson writes that we should allow “sharers of a scheme” to “differ on the details” (VICS, 194).) Davidson writes that this is “a useful criterion” (VICS, 193) for differences in conceptual scheme only if we can understand “the notion of truth, as applied to language, independent of the notion of translation” (VICS, 193). He argues that we cannot: Tarski’s Convention T is “our best intuition as to how the concept of truth is used” (VICS, 193), and instances of Convention T are translations of an object
language into a meta-language. Thus, we understand the concept of truth as applied to language only via “translation into a language we know” (VICS, 193).

Davidson concludes from this that the present formulation of conceptual relativism cannot provide “a test that a conceptual scheme is radically different from ours”, as it “depends on the [false] assumption that we can divorce the notion of truth from that of translation” (VICS, 193). If we are to suppose that something is a (largely) true theory, we must also suppose that it is, at least in principle, translatable into “a language we know”. But then the theory in question cannot be identified or associated with an alternative conceptual scheme, since an alternative scheme (if such there be) must be associated with a language or theory that cannot be translated into our language.

This complex line of argument depends on two attempted conceptual reductions. First, Davidson claims that the concept of fitting the totality of sensory experience or evidence reduces to the concept of truth. The first concept “adds nothing intelligible” (VICS, 193-4) to the second: its content is nothing more than that of the concept of truth. Then he claims that the concept of an entity that “makes” sentences or other vehicles of truth true (or false) reduces to the “simple” concept of truth. So Davidson here claims that a certain traditional theory of truth, which might seem to encourage the scheme/content distinction, is at best a misleading expression of a “simple” concept that lends no support to the distinction. (In a footnote, Davidson refers us to his essay “True to the Facts”, where an important argument against traditional correspondence theories of truth is set out.95 This argument is discussed in some detail in the next chapter.)

Thus, neither the concept of sensory evidence nor the correspondence theory of truth will allow us to “add a new entity to the universe against which to test conceptual schemes” (VICS, 194). This does not establish that the very idea of a difference in conceptual scheme, on the present account, is unintelligible. At most, what has been established is that this idea reduces to the “simple thought” of a true but untranslatable theory. Although this suggestion is not obviously intelligible, it is not obviously unintelligible either. So the final phase of Davidson’s argument, as we’ve seen, is intended to show that the concepts of truth and translation are inter-dependent, with the

95 Donald Davidson, “True to the Facts”, in Inquiries into Truth and Interpretation, 42.
consequence that this “simple thought” is in fact unintelligible. We’ll consider this final phase of the argument in the next chapter.

II

There is a great deal to say about the argument outlined in the last section. In this chapter, I will focus almost entirely on Davidson’s subordinate argument for the assimilation of the concept of fitting the totality of sensory experience or evidence to the concept of truth. First, however, we should consider how exactly the argument outlined above is supposed to contribute to Davidson’s master argument against the intelligibility of the very idea of a conceptual scheme.

The conclusion of the master argument, recall, was supposed to be that all criteria of identity for languages, and hence conceptual schemes, imply or depend on the possibility of translation into our language. This would show that “a form of behavior that is not interpretable as language in our language is not speech behavior” (VICS, 184). Davidson promises an argument for this thesis, which he concedes is far from obvious. The argument, seemingly, proceeds by examining the two main proposals for a criterion of identity for languages that does not imply or depend on the possibility of translation into our language, and by showing that these proposals fail.

If Davidson’s discussion here is to support his main thesis that the idea of a conceptual scheme is unintelligible, it should provide a reason to reject the metaphor of “fit” as a basis for a criterion of identity for languages (or theories) independent of the possibility of translation into our language. The conclusion of the argument under discussion, however, seems to be merely that there is no intelligible criterion of identification for alternative conceptual schemes—no “test”, or “useful criterion” for identifying differences in conceptual scheme. (A useful criterion, presumably, is one that we can apply, one that we can use to test for differences in conceptual scheme.) Be that as it may, it does not follow from this that there is no intelligible criterion of identity for conceptual schemes. Certainly it does not follow that there could not be alternative conceptual schemes—not, at any rate, unless we are to assume from the outset that an
intelligible criterion of identity must somehow imply the possibility of identifying alternative schemes.

If we were to grant Davidson’s view that to identify some form of behavior as language is to recognize it as something susceptible to translation or interpretation, we could assume that “a form of behavior that is not interpretable as language in our language is not speech behavior” (VICS, 184). But this was the very conclusion that Davidson’s argument was supposed to establish. On the other hand, if Davidson means only to argue here that we cannot identify alternative conceptual schemes, i.e., that they are “not interpretable as language in our language”, the present conclusion seems irrelevant to his master argument. Given the way he has framed the larger issue, the present conclusion reduces to the tautology that if there are alternative conceptual schemes (i.e., languages that are “not interpretable as language in our language”) they are alternatives to our conceptual scheme. So Davidson’s conclusion seems either question begging or irrelevant to his main line of argument.

In what follows, I’ll assume that Davidson’s argument is for the conclusion that the metaphor of “fit” fails to provide a criterion of identity for alternative conceptual schemes, and not that it fails to provide a criterion of identification. As he says earlier in “On the Very Idea of a Conceptual Scheme”, “the problem” facing the conceptual relativist “is to say what the [scheme/content] relation is, and to be clearer about the entities related” (VICS, 191). So it is not, or not necessarily, incumbent on the conceptual relativist to say, in addition, how we might identify alternative conceptual schemes, how we might provide a “test” or “useful criterion” for differences in conceptual scheme. The onus is on the conceptual relativist to give some intelligible account of how schemes are related to some other entity, and what that entity is. But the account need not establish that, if there were alternative schemes, we would be able to identify them.

We should also note that while the present account of conceptual schemes is based in the relation of theory to evidence, Davidson earlier associated conceptual schemes with languages. While he denies that this account of conceptual schemes is intelligible, he seems untroubled by the assimilation of language to theory. Does it make any sense, however, to suppose that, like a theory, a language may be “borne out by the
evidence”, or “largely true”? Certainly sentences or groups of sentences, theories, may be borne out by the evidence, or may serve to predict future experience; and they may be true (or “largely true”). Obviously, though, we can’t infer from this that the language in which those sentences are formed—say, a certain grammar and vocabulary—is the kind of thing that predicts anything, that can be “borne out by the evidence”, warranted, true or false, etc. This would be like reasoning that, since some English sentences express questions, the English language is itself a question.

The rejection of the analytic/synthetic distinction is supposed to motivate the assimilation of language to theory: “to give up the analytic-synthetic distinction as basic to the understanding of language is to give up the idea that we can clearly distinguish between theory and language” (VICS, 187). But even on the debatable assumption that we cannot clearly distinguish between theory and language, it does not follow that we cannot distinguish between theory and language at all, or that there is no distinction. There is no clear distinction between young and old, but it doesn’t follow from this that everyone is old, or that there is no way for us to distinguish between young people and old people.

Just as importantly, the upshot of Quine’s attacks on the analytic/synthetic distinction is supposed to be that there is no sharp distinction between purportedly analytic and purportedly synthetic sentences of a language: no sharp distinction between sentences or theories true in virtue solely of linguistic convention and those true in virtue of linguistic convention and extra-linguistic reality. Even if no sentence is true solely in virtue of linguistic convention, this would not explain how languages, as opposed to (declarative) sentences belonging to those languages, could be seen as true or false, warranted or warranted.

Presumably Davidson is here alluding to the idea that the linguistic conventions of a language commit speakers of that language to some very general conception of reality, in that these conventions are or are associated with some system of basic categories. For example, we noted in Chapter 1 that, according to Quine, our beliefs about enduring material objects comprise just one among countless empirically equivalent theories of the world. Insofar as the grammar and vocabulary of a language contain devices for individuating objects, then, we might think of the language itself as a kind of theory. Or,
more plausibly, there is Quine’s view that even the most basic observational terms of a language are, in some sense, theory-laden. That is, no system of categories can be derived from sensory experience or stimulation alone. It’s hard to make sense of Quine’s suggestion that there could be a theory without individuation. A more intelligible example might be to suppose that the linguistic conventions governing the genders of nouns in French or German commits speakers to a “theory” of the genders of inanimate objects. Despite the ostensible shift in metaphors, it would seem that we are concerned, again, with the notion of conceptual or linguistic organization.

I take it that this is the general kind of position that Davidson has in mind. However, notice that, even on this kind of view, it still does not make much sense to say that a language is “borne out by the evidence”. For how could the use of the basic observational terms in a language, on this view, come into conflict with sensory experience or evidence? Of course, we might decide that some particular utterance of “That’s red”, for example, was false. The present suggestion, by contrast, is that there is some sense in which merely to make use of our language, with its observational terms, is somehow to commit to a certain theory of the world that is warranted by the sensory evidence. Thus, a later judgment that the utterance was false, and that (say) the thing indicated is actually blue, would be just another move within the same overall theory of the world.

The problem is that, on this kind of view, all the evidence available to speakers of a language is theory-laden. We could not find evidence against our language, scheme or theory because that language determines what evidence there is: it is (partly) constitutive of the evidence. To be sure, a theory may be “borne out by the evidence” in a vacuous sense if the evidence (or the description of the evidence) somehow presupposes the truth of the theory. But then it would be more accurate to say that the theory is “borne out” by itself, i.e., that the truth of the theory implies the truth of theory. So, while this conception of language may suggest a way to make sense of the claim that languages are true, it is not clear whether it suggests a way to make sense of the claim that they are “borne out by the evidence”. For if it is not possible even in principle to find evidence against some theory, it is not clear how there could be evidence for the theory.
In any case, let us assume for now that there is some sense in which languages can be warranted or “borne out” by the evidence and, also, some sense in which they may be “largely true”. (To say that a language (qua theory) is “largely true”, presumably, is to say that the organizing principles or categories associated with the basic observational terms of the language somehow reflect or correspond to the basic features of reality as it is in itself. For example, if a belief in the reality of distinct and “enduring physical object[s]” is a basic part of our conceptual scheme or theory, the scheme or theory is true, in that respect, if there really are distinct and enduring physical objects.)

In light of these considerations, we can motivate Davidson’s assumption that “what is in view here is not just actually available evidence” but rather “the totality of possible sensory evidence”. The present suggestion is that “sensory evidence” is something like “sensory experience” or “the pattern of our surface irritations”, i.e., some form of pre- or non-conceptualized sensory input. But this alone—some sensory experience, or the occurrence of a certain sensory experience—could not warrant the acceptance or rejection of sentences or theories. Rather, sensory experience serves as the evidence for sentences or theories only insofar as it is “organized” or “systematized” by means of some language or theory. Strictly speaking, only reports of sensory experience (e.g., Quinean observation sentences) can constitute evidence for or against our theories and beliefs.

So, again, it makes no sense to suppose that any evidence could count against a conceptual scheme, as any putative evidence against a conceptual scheme would itself be framed in terms of some scheme or other: the evidence is some body of observation sentences or reports, and therefore belongs to some particular language or theory. Hence these observations must either belong to a language that can be translated into the language(s) associated with the scheme in question, or to a language that cannot be translated into the language(s) associated with that scheme. (In other words, the evidence is either interpreted or “organized” according to the scheme under discussion, or according to some other scheme.) But in neither case can the evidence, so understood, come into conflict with the conceptual scheme.

If the evidence is a body of sentences that are translatable into the language(s) associated with the conceptual scheme in question, to count those sentences true would
be to accept the conceptual scheme to which they belong. But if, instead, the evidence is a body of sentences that are not translatable, even in principle, into the language(s) associated with that conceptual scheme, they cannot stand in rational relations to the scheme. More intuitively, the point is that the mechanism by means of which we acquire evidence could not intelligibly be called into question by the evidence thus acquired.

Thus, it seems that any attempt to find evidence against a conceptual scheme will be self-defeating. If there is some (actually) possible observation or piece of evidence that favors one theory over the other, then presumably both can be expressed and compared by means of a third theory; however, this “belies the claim of dramatic incomparability” (VICS, 184). For in this case, T1 implies that $s$ while T2 implies that not-$s$; presumably, then, T1 and T2 contain a significant range of inter-translatable terms. But if so, it is not clear why we should consider this a difference in scheme rather than a difference of opinion arising within a single scheme. Therefore, it seems we should conclude that any one conceptual scheme is empirically ideal; there is no evidence that could count against it. But, as we noted earlier, this may also be a reason to doubt whether it is intelligible to say that a conceptual scheme is “borne out by the evidence”.

Finally, recall that, according to Davidson, there are at least two kind of “entities that get organized, or which the scheme must fit”: “either it is reality (the universe, the world, nature), or it is experience (the passing show, surface irritations, sensory promptings, sense-data, the given)” (VICS, 192). But as we’ve seen, Davidson’s

---

96 Maybe I should clarify this point. What may be unintelligible, it seems to me, is the idea that a theory might be ideally (empirically) “warranted”, if the reason for this is that any evidence available to proponents of the theory is framed in terms of that same theory. But I don’t mean to suggest that the very idea of an ideally warranted theory is unintelligible. For perhaps some theory is ideally warranted because it is true: any true theory, presumably, is warranted by all (actually) possible evidence, i.e., the totality of evidence that any actual observer could have. It’s doubtful whether the first kind of ideal warrant is intelligible because its ideal status is due to the epistemic limitations of proponents of the theory. Since they aren’t able to imagine that the theory is false, it’s not clear that they can be said to believe that, given the evidence, the theory is true. (And for that reason, perhaps it’s also unclear whether they can be called “proponents of the theory”, or whether the “theory” really is a theory.) The second kind of ideal warrant is due, instead, to the metaphysical relation of the theory to reality. There couldn’t be evidence against a true theory simply because it is true, and what evidence there is—as opposed to what might seem to be evidence, that is—depends on what is true. (Of course, there might seem to be evidence against it, if those evaluating the evidence are less than omniscient.) In this latter case there is no problem in imagining that proponents of the theory believe, on the basis of whatever evidence they have, that the theory is true. For we are not assuming that any evidence they could discover would presuppose the truth of the theory. Therefore, it may be that they can imagine evidence against the theory, and hence imagine that the theory is false. As it happens, though, they never do or could find any such evidence.
discussion of the metaphor of “fit” seems to focus exclusively on the second suggestion. The metaphor of “fit”, he thinks, is based in the “general position” that “sensory experience provides all the evidence for the acceptance of sentences (where sentences may include whole theories)”. This is not, however, what it would mean to say that a sentence or theory fits reality, the universe, the world or nature. This may have nothing much to do with the notion of evidence or justification.

Why, then, does Davidson not explore the proposal that languages, theories or conceptual schemes fit reality (or the world, or nature)? He does mention this idea, if only in passing: “The trouble is that the notion of fitting the totality of experience, like the notion of fitting the facts, or of being true to the facts, adds nothing intelligible to the simple concept of being true” (VICS, 192). His point seems to be that all of these seemingly different ways of elaborating on the metaphor of “fit” ultimately reduce to the claim that conceptual schemes (qua theories) are true. But it is not obvious that “the notion of fitting the totality of experience” reduces to the concept of truth, so Davidson makes some attempt to explain this. Whether the entity that schemes are supposed to “fit” is sensory experience or evidence or, instead, reality or “the facts”, the conclusion, he thinks, is that conceptual schemes are true.

Neither is it obvious, however, that “the notion of fitting the facts … adds nothing intelligible to the simple concept of being true”. This would seem to be the idea of a correspondence between statements, propositions, beliefs, sentences or other truth-valuable entities, on the one hand, and the objective reality in virtue of which they are true or false, on the other. Thus, we might suppose that the notion of fitting or corresponding to the facts does “add” something to “the simple concept of being” true. Perhaps it “adds” an explanation of the nature of truth.

However, Davidson holds that correspondence theories of truth are untenable, for reasons we will consider later. Whether or not his objection to correspondence theories of truth is sound, a theory that fits or corresponds to reality must in any case be true; that is, whether the claim that the theory “corresponds” to reality explains why it is true, or just means that it is true, it must be true. Hence, in either case, this line of thought may be vulnerable to Davidson’s objection that we cannot make sense of the idea of a true but untranslatable theory. If his argument against correspondence theories is not sound,
however, the objection presently under discussion will require an additional premise to
the effect that truth \textit{qua} correspondence implies that translation, or, at least,
interpretation, is possible across languages that correspond to the same facts.

In a later chapter, I’ll argue for this claim. For the moment, we can set aside the
question of whether correspondence theories of truth are defensible, which will be
considered in the next chapter. But it is worth noting his complication in Davidson’s
reasoning, since it would otherwise be puzzling that he does not explore the idea that
fitting “the facts” or “the world” provides a basis for the scheme/content distinction quite
different from the idea of fitting experience or evidence.

III

To summarize the main points of the last section, we’ve found that at least some
of the initially puzzling features of Davidson’s discussion can be clarified and defended:
(i) Davidson’s argument should be taken to concern the possibility of a criterion of
identity, not a criterion of identification, for conceptual schemes;
(ii) There may be a way to make sense of the idea that languages are “borne out by the
evidence”, and, perhaps, the idea that they are “largely true”;
(iii) Conceptual schemes, on the present account, are empirically ideal, i.e., “borne out”
by “the totality of possible sensory experience”; (iv) However exactly the notion of fitting reality, the world or the facts is to be
understood, it implies truth; therefore, we do not need to consider this alternative to the
notion of fitting evidence or experience.

What remains unclear is the connection between the “general position” Davidson
describes and the thesis that differences in conceptual scheme are logically possible, i.e.,
that there could be two theories or languages that are not inter-translatable (in the relaxed
sense explained and defended earlier). Neither is it entirely clear what that “general
position” comprises. Davidson’s remarks don’t establish whether the association
between ideal empirical warrant and truth is a basic feature of that position, or one of its
consequences, or (less plausibly) something that Davidson himself believes and takes to
be obvious enough to require no argument. (Davidson writes simply: ‘the point is that
for a theory to fit or face up to the totality of possible sensory evidence is for that theory to be true” (VICS, 193). It isn’t clear to which point he is referring.)

One reading is that the “general position” Davidson describes is simply a kind of knowledge empiricism: the position is that sensory stimulation or experience (*qua* evidence) is the only justification for our beliefs (or for “the acceptance of sentences”). The account of conceptual schemes that supposedly stems from this position is that there could be two or more empirically ideal theories. Given the poverty of the sensory evidence, there could be two or more theories that are “borne out” by “the totality of possible sensory evidence”.

The “general position”, then, is or implies something like Quine’s radical version of the under-determination thesis. As Quine puts it, this is the view that we have no reason to suppose that man’s surface irritations even unto eternity admit of any one systematization that is scientifically better or simpler than all possible others. It seems likelier … that countless alternative theories would be tied for first place.97

But the most natural conclusion to draw from this would surely be that, even in the ideal case, there could be no way to tell which of “countless” incompatible theories is true. This would be an interesting and unsettling conclusion, but it has no bearing on the thesis that alternative conceptual schemes are possible, as the fact that some other theory is empirically equivalent to our own does not preclude the possibility of expressing both theories in a single language.

Indeed, this conclusion is inconsistent with the assumption that two theories “tied for first place” in this way are associated with distinct conceptual schemes. For if there would be, in this situation, something that it is impossible for us to decide—the question of which of the two or more theories is true—these various empirically ideal theories must be in cognitive disagreement, each implying the negation of the other. Otherwise, we could view each as complementing each or every other theory, describing aspects of the universe that the other(s) do not. But then there is no sense in which the theories are “tied for first place”; they are not in competition to begin with. So, it appears that the theories must be in cognitive disagreement; they must be somehow inconsistent. And

yet, as we noted earlier, cognitive disagreement or inconsistency across conceptual schemes is impossible.

On another reading, it may be that the “general position” Davidson describes comprises not only knowledge empiricism, and a skeptical view of the evidence available to us, but also an epistemic theory of truth, according to which truth is ideal empirical warrant. But even so, it remains unclear how this line of thought might encourage conceptual relativism. If two ideally warranted theories are both true, they might be complementary portions of some larger theory. Or, instead, they might be synonymous (i.e., they might be merely notational variants on a single theory). In neither case would we have reason to suppose that they would not be inter-translatable. To be sure, the present line of thought appears to provide a conception of language that does not explicitly imply or presuppose the possibility of translation into our language. But it also seems quite likely that this is nonetheless implied or presupposed, as the preceding discussion suggests. So it is unclear how it could be even superficially reasonable to hold that two ideally warranted theories might be associated with different conceptual schemes. At the very least, we need some hint as to why two theories supported by exactly the same evidence would be associated with incommensurable concepts.

In Chapter 1, we explored in some detail how Quine argues, at times, for conceptual relativism from the under-determination thesis. However, Quine’s argument typically hinges on the idea of a conceptual or linguistic organization of sensory stimulation or experience, “systematizations” of “man’s surface irritations”. This was suggested earlier, as a way of making sense of the claim that a language might be intelligibly described as something “borne out by the evidence” or “largely true”. Because these “countless” different systems of organizing principles are radically under-determined by the evidence, the basic observational terms or categories of one language will most likely have no counterparts in another. There may be no overlap between the sets of stimulatory episodes picked out by the observational terms of any two languages.

Two such languages or theories would perhaps both be “largely true” even though neither is translatable into the other. However, on this view, the failure of translation between languages or theories is not explained by appeal to the “general position” Davidson describes, but, instead, by appeal to “the referential apparatus of language”.

132
But then the present formulation of conceptual relativism reduces to the first, and is vulnerable to the objection discussed in the previous chapter. To repeat, the objection is that such failures of translation are intelligible, but only on the assumption that both of the languages or theories under discussion are referentially related to a single ontology. However, the ontology common to both languages or theories provides a basis, in principle, for interpreting both of these by means of some third, more comprehensive language or theory. This possibility is at odds with the claim that the two original languages or theories are associated with divergent conceptual schemes.

IV

I propose that we interpret Davidson’s argument roughly as follows:
(1) By hypothesis, alternative conceptual schemes are associated with languages or theories that are not translatable into our language.
(2) The thesis that conceptual schemes “fit” some entity implies that alternative conceptual schemes are (largely) true.
Therefore,
(3) On the present account, an alternative scheme is (or is associated with) a language or theory that is (largely) true but not translatable into our language.
However,
(4) The concept of truth is not intelligible independently of the concept of translation.
Therefore,
(5) To suppose that some theory or language is true is to suppose that it is translatable into our language.
Therefore,
(6) On the present account, a supposedly alternative conceptual scheme is (or is associated with) a language or theory that is translatable into our language.
Therefore,
(7) The present account fails to establish that there could be differences in conceptual scheme.
As we’ve seen, the “general position” Davidson describes could be taken to provide an account of differences in conceptual scheme, independently of the idea that languages, theories or schemes organize experience, only if such an account can be based merely on the concept of ideal empirical warrant and, perhaps, an epistemic theory of truth. These alone do not suggest any clear reason to suppose that two empirically ideal theories would not be inter-translatable. On the contrary, purportedly different schemes, on this kind of view, would seem to be rival theories arising within a single language or scheme. But let us suppose that some such account might be devised, and consider the argument outlined above.

The argument is perhaps best seen as a *reductio*: premise (1) asserts that alternative conceptual schemes are not translatable into our language, while the subordinate conclusion (5) establishes that they are translatable into our language. Hence alternative conceptual schemes, on the present account, are both translatable and non-translatable into our language. Alternatively, the conclusion of the *reductio* might be that alternative schemes are both true and not true. Premise (2) asserts that alternative schemes are true, but from (1) and (5) we can infer that they are not true.

In any case, the important point is that, according to Davidson, the present account of conceptual schemes is incoherent because of the relations between the concept of the totality of possible sensory experience or evidence, the concept of truth and the concept of translation. The most questionable premises in this reconstruction of Davidson’s argument are (2) and (4).

In the next sections, I’ll consider some of the issues surrounding premise (2), according to which the notion of evidential “fit” somehow implies truth. (As we’ve seen, Davidson says that this just *is* the notion of truth, but for now we can consider the weaker claim that truth is implied by talk of “fitting”.) First, let us consider the *prima facie* plausibility, apart from its role in Davidson’s discussion or, as may be, in the general position he takes to encourage conceptual relativism. Certainly talk of “fitting” the facts, reality or the world would seem to imply truth, as we’ve seen, but what of “the notion of fitting the totality of experience” (VICS, 193-4)?

As Davidson notes, this is the idea of a language, theory or conceptual scheme that is ideally *warranted* by experience or evidence. This would seem to be an altogether
different notion, with no obvious connection to the concept of truth. And yet Davidson writes: “if a theory quantifies over physical objects, numbers, or sets, what it says about these entities is true provided the theory as a whole fits the sensory evidence” (VICS, 193). This claim is not credible if the “sensory evidence” is evidence that we could have, for reasons considered earlier. Seemingly, a theory might be warranted by all the evidence available, in principle, to us and yet false—or at least, less than “largely true”. For example, it seems possible that there are many fundamental features of the universe that we aren’t able to perceive or conceptualize; we might not have the right sense organs, or the right kind of brain. If our theory of the universe purports to describe the basic laws or structure of the universe, then in this case it would, arguably, be less than “largely true”.

So it is far from obvious that a theory supported by all the evidence we could have could not also be false, or less than largely true, and some argument is needed for that claim. However, Davidson simply asserts argument that the concept of “fitting” the totality of possible evidence reduces to the concept of truth. It would be more defensible to claim that a theory warranted by all possible evidence must be largely true if “the totality of possible sensory evidence” refers instead to the totality of observations made by all logically possible observers. (Of course, this cannot mean the totality of observations made by observers across all possible worlds, since no theory could be warranted by that set of observations. What I have in mind here is instead the totality of observations, evidence or experience had by all the observers that there could have been, in principle, in the actual world.) Certainly if an omniscient observer is logically possible, a theory that is ideally warranted, in this sense, must be true. (We can set aside the question of what it would mean for an omniscient being to have “evidence”.)

On this interpretation, Davidson is clearly right to say that nothing has been “added” to the concept of truth. The idea of a theory warranted by all possible evidence, in this sense, is just the idea of a true theory. But this is plausible only because the reference to evidence or experience is superfluous. We have no real idea of what the relevant totality of evidence or experience is supposed to be apart from its connection with the concept of truth: it is just the body of evidence or experiences that would be had by an observer whose beliefs were all (or largely) true. Unless we suppose that there are
certain facts or features of reality that are unobservable in principle, this notion of evidence thus collapses to the notion of fact, reality or truth.

However, this vacuous claim is surely not what Quine and other proponents of the scheme idea have in mind when they speak of “experience” or “evidence”. They intend to discuss the actual beliefs and theories of human beings (among other logically possible beliefs and theories, perhaps). At any rate, they presumably intend to discuss the beliefs and theories of non-omniscient beings—as in Davidson’s brief discussion of Saturnians and Plutonians, for example (VICS, 186). But we can’t explain our actual use of language or theory construction by reference to a body of evidence that is not available, even in principle, to being like ourselves. On this interpretation, then, the general position under discussion is not a kind of knowledge empiricism; it reduces to the tautology that true theories cannot be false. But this would be a radically uncharitable construal of that position, which purports to say something substantive about conceptualization and theory construction.

Davidson’s reasoning in this passage thus faces a dilemma. On the one hand, the evidence available to any less than omniscient being, even the totality of evidence available to the totality of logically possible non-omniscient beings, is not sufficient to imply the truth of any theory warranted by that body of evidence. There is always the logical possibility that the totality of facts about universe exceed the perceptual and cognitive faculties of any and all such beings. But if this is possible, then insofar as any theory that can be devised by a less than omniscient being purports to describe the basic, general features of the universe, it is doubtful whether any such theory must be considered “largely true”.

On the other hand, if the evidence in question is identified instead by an implicit appeal to the totality of truths, the general position he describes says nothing substantive about our theories or concepts. Now, if Davidson’s reasoning is sound, it may be that that kind of position, which is vacuously true, fails to provide an account of conceptual schemes. But his argument does not show that no such account is possible on the basis of some more substantive position concerning the relation of theory to evidence or experience. So it would seem that premise (2) in my reconstruction of Davidson’s
argument is either very implausible or else irrelevant to the question of whether, as Quine and others have suggested, knowledge empiricism encourages conceptual relativism.

This reflects a familiar and more basic problem with epistemic theories of truth. If talk of fitting all possible “evidence” or “experience” is construed so as to imply truth, these seemingly epistemic terms lose all epistemic content. But if we take these to be genuinely epistemic terms, and not merely synonyms for “true”, the logical connection with truth is dissolved. As Davidson says elsewhere:

Either the conditions of warranted assertibility [i.e., the conditions that are supposed to be conditions for truth] are made so strong that they include truth itself, in which case the account is circular, or circularity is avoided by making the conditions explicit, and it then becomes clear that a fully warranted assertion could be false.98

Thus, if the present suggestion is that conceptual schemes (qua theories) are warranted by the totality of evidence available to an omniscient observer, or to all logically possible observers, the identification of warranted belief, in this absurdly stringent sense, can be plausibly identified with truth. But clearly, on this reading, “the conditions of warranted assertibility are made so strong that they include truth itself”. But if, instead, we suppose that conceptual schemes (qua theories) are warranted by some lesser body of evidence, “it becomes clear that a fully warranted assertion [or theory] could be false”.

I agree with Davidson that epistemic theories are vulnerable to this dilemma. If the theory sets out conditions necessary and sufficient for truth, these must implicitly appeal to some prior, non-epistemic notion of truth. If the theory sets out conditions that do not ultimately depend on a prior, non-epistemic notion of truth (i.e., genuinely epistemic conditions), it will turn out that they are neither necessary nor sufficient for truth. But why then does Davidson seem to assume, in his discussion of the metaphor of “fit”, that only the first horn of the dilemma is possible on the present account of conceptual schemes? Isn’t it possible, as I’ve been urging, to interpret talk of “fitting the totality of experience [or evidence]” in a genuinely epistemic way, with the consequence that a theory could be warranted by that evidence and yet false?

On the face of it, Davidson’s contention that the idea of a theory that fits the totality of experience or evidence reduces to the idea of a true theory is not credible. Nor does it seem a fair interpretation of the “general position” under discussion—especially if that is supposed to be Quine’s position, which comprises a radical version of the under-determination thesis. If Davidson’s argument is to apply to this general kind of view, and not merely to a vacuous position unrelated to knowledge empiricism, premise (2) must be supported by a subordinate argument. The subordinate argument will have to justify either the conclusion that, despite the objections pressed in this section, there is some non-vacuous conception of ideal empirical warrant that implies truth, or the conclusion that some such epistemic theory of truth is essential to the general position under discussion.

V

Ernest Lepore and Kirk Ludwig argue that Davidson’s introduction of premise (2) appeals to a kind of meaning empiricism, according to which terms of a theory are introduced explicitly in the context of giving an account of the evidence the theory is meant to accommodate… On this view, if a theory accommodates all the possible evidence, there is no additional question about whether the world is the way the theory says it is. This would presuppose that there were some fact of the matter about what the theoretical terms in the theory meant beyond their contribution to accounting for possible evidence for the theory. This view, applied to interpretation theory, underlies Davidson’s thesis of indeterminacy of interpretation.99 On this (broadly) empiricist theory of meaning, the meaning of theoretical terms is constrained by the circumstances in which they are introduced. It is supposed to follow from this that a theory could not be ideally warranted and yet false. Essentially, the idea seems to be that a theory “says” nothing about the world as it is independently of all possible evidence. The truth conditions of the theory concern only “the evidence the

theory is meant to accommodate”. Since the evidential facts exhaust the meaning of the theory, there is “no additional question about whether the world is the way the theory says it is”. So the theory is, equally, an epistemic theory of truth, which might justify premise (2). This would seem to be the view that underwrites Davidson’s claim that a theory is true if (as a whole) it “fits” the sensory evidence.

My main interest here is whether this theory of meaning and truth provides a good subordinate argument for premise (2)—i.e., the premise that ideal empirical warrant implies truth. However, it is worth noting, in passing, that it is very unlikely that Davidson actually holds the theory of meaning that Lepore and Ludwig describe. Essentially, this would seem to be “the theory that talk about brick houses on Elm Street is ultimately to be construed as being about sense data or perceptions”, or something very much like it, which Davidson calls an “extreme and implausible” (VICS, 193) version of the position under discussion.

Elsewhere, Davidson writes, “I have always been clear that I was not an antirealist about any theoretical entities over which a theory I held true quantified”.\(^{100}\) This reflects his acceptance of Quine’s criterion of ontological commitment. If we hold that a quantified sentence containing a predicate F is true, we are committed to the existence of some thing or things of which F is true unless we can provide a paraphrase of the sentence that does not contain F. According to Quine, it is incoherent to assent to the sentence “Molecules are made of atoms”, for example, while denying that molecules are real, unless one is able to re-express that purported truth by some sentence that does not contain the word “molecule”. In any case, as we’ve seen, Davidson explicitly rejects epistemic theories of truth.

But if Davidson is, in this sense, a realist (or, at least, not an anti-realist) about the entities over which theories he holds true quantify, he cannot also hold that the meaning of a term like “molecule” is somehow exhausted by its role in accommodating evidence for the theory. For presumably molecules are not among the evidence for any theory we could devise. Thus, on the empiricist theory of meaning under discussion, it would seem

---

that whatever entities are the values of the variable in sentence “∃x (molecule (x))”, none of these are molecules. Obviously this conclusion is anti-realist.

To be sure, Davidson’s view of “interpretation theory” is as Lepore and Ludwig say. With Quine, Davidson claims that theories of meaning and belief are indeterminate; but surely it is significant that neither argues for the indeterminacy of physics (or chemistry, or biology, or any other scientific theory). Rather, they hold that physical theories (and others), unlike theories of meaning and belief, are under-determined by all possible evidence. This is inconsistent with the theory of meaning that Lepore and Ludwig ascribe to Davidson. Suffice to say, then, that this ascription seems very implausible for a number of reasons.

In any case, let us consider whether this (radical) empiricist theory of meaning and truth provides a good subordinate argument for premise (2). Again, this is the premise to the effect that ideal empirical warrant implies (or perhaps just is) truth. If it does provide a good argument for that conclusion, then perhaps Davidson’s reasoning can be defended on grounds other than those he gives (and other than those he would accept). I think it is fairly easy to show that this theory fails to provide a good argument for premise (2), because it is no more warranted than the bare assertion of premise (2). We can see this by reviewing the difference between under-determination and indeterminacy.

A theory is under-determined with respect to certain evidence if there is some other, incompatible theory would account for the same evidence equally well. It seems to be conceivable two theories be equivalent and yet incompatible with respect to all possible evidence. However, this would not necessarily be a case of indeterminacy, for it might be that one is true and the other is false. Indeterminacy requires, in addition, that there is no fact of the matter as to what the competing theories say over and above their

---

101 As we’ve seen earlier, there is nevertheless a serious tension between the thesis of the indeterminacy of translation or interpretation and the realist conception of physics that Quine and Davidson share. The indeterminacy of interpretation and linguistic meaning generally infects what we say and think about physics as much as any other subject matter. But it is hard to imagine how any facts about observable behavior in public circumstances could determine that a speaker is referring to molecules or the Big Bang. Therefore, it may be that the indeterminacy of translation or interpretation implies the indeterminacy of physics (or at least, of our physical discourse), collapsing the distinction between theories of meaning and belief and other kinds of theories. However, Quine and Davidson clearly think the distinction can be maintained.
“contribution to accounting for possible evidence for the theory”. (This is Davidson’s notion of indeterminacy at least, which differs from Quine’s in that, as we’ve seen, Quine is often inclined to hold that the competing theories do “say” something more than this. But we can ignore these complications)

Lepore and Ludwig give the example of (seemingly) rival decision-theoretic accounts of belief and desire:

In decision theory, … the concepts of degree of belief and desirability … receive empirical interpretations in terms of an agent’s choice behavior, that is, the pattern of the agent’s choice behavior wholly determines what his degrees of belief and desirability are: there can be no further hidden facts about them. 102

If two such theories of an agent’s “degree of belief and desirability” are empirically equivalent, there is no further, intelligible question as to “whether the world is the way [one or the other] theory says it is”. This is quite plausible as concerns the concept of a degree of belief and desirability. Perhaps it is less so as concerns the ordinary, concept of belief or desire, which is not acquired through an understanding of decision theory. But a case can be made that the latter concepts are modeled fairly accurately by decision theory, as Davidson maintains, and that all understanding of linguistic meaning and mental life is similarly indeterminate as between empirically equivalent accounts.

But is it reasonable to hold that the seemingly distinct domains of evidence and fact collapse in the same way for most or even many other kinds of theories—that, for instance, the totality of historical evidence “wholly determines” what happened in the past? According to premise (2), to say that a theory is borne out by all possible evidence or experience is just to say that the theory is true. No distinction is made between kinds of theories. Moreover, one of the theories under discussion is our own overall theory of the world, i.e., our own “conceptual scheme”, which comprises history, physics and much else. So if the argument from meaning empiricism that Lepore and Ludwig present is to support premise (2), it must have as a conclusion the claim that for any theory, and not only decision theoretic accounts of belief and desire, the totality of evidence wholly determines the relevant domain of facts.

102 Davidson, 224.
But as applied to historical or physical theories, for example, this thesis does not seem to be true. Given what we believe about the objects of these other kinds of theories, it seems entirely possible that there are “hidden facts” that are not discoverable on the basis of the kind of evidence we could obtain. Even given all possible evidence that we could realistically imagine acquiring, it would seem to be logically possible that some or even most of our beliefs about these matters would still be false—beliefs about Charlemagne’s diet or the details of Aztec cosmology, about the behavior of subatomic particles or the origin of the universe, and so on.

My point here is that it is in the nature of these matters, as we normally conceive of them, at least, that no amount of evidence or experience available to us could wholly determine the relevant facts. Unlike the controversial case of theories of meaning and belief, we normally take the truth of a historical or physical theory is determined by the historical or physical facts, which are logically independent of the historical or physical evidence. For example, suppose that a historian claims that there were 500 inhabitants of a certain Aztec village in 900 A.D. The literal meaning of this claim surely appears, at least, to be about nothing more or less than the number of people in that village at that time. It does not appear to say or imply anything about the evidence or experiences available to historians.

So it seems to me that there is a strong presumption in favor of a realist view as concerns most theories. Some argument is needed to show that, in some particular case, the distinction between warranted belief and truth collapses. Now there is an argument of this kind as concerns theories of linguistic meaning and mental life. As we’ve seen, Davidson and Quine make much of the fact that we learned to speak only by observing the behavior of others in shared, public circumstances. Therefore, the facts of linguistic meaning and, by extension, mental life generally could not radically transcend all the evidence available to us. (Again, it is worth emphasizing that here we are speaking of the languages and minds of creatures sufficiently similar to us, i.e., creatures that could be part of our public. It does not follow from these claims about language learning and meaning that there could not be radically alien publics, which we would be unable to penetrate, and so perhaps alternative conceptual schemes.)
These are thus special considerations in favor of the conclusion that if a speaker means or believes that \( p \), this is something we must, in principle, be able to discover, and that if there is no evidence that can settle whether he believes that \( p \) or, instead, that \( q \), there can be no “hidden facts” here. It could not be that he in fact believes that \( p \), and does not believe that \( q \), even though no human observer could ever discover that these mental facts. By contrast, there are no such special considerations in the case of physics (or biology, history, etc.). So far as we know, there is nothing about the behavior of subatomic particles, say, or events in the distant past that would imply that the relevant facts are by their very nature accessible to us.

VI

This important asymmetry between theories of meaning and belief and other kinds of theories, together with the presumption in favor of a distinction between warranted belief and truth, is probably a sufficient reason to reject the form of meaning empiricism that Lepore and Ludwig ascribe to Davidson. But there are other good reasons. I’ll mention just two more, both of which concern the seemingly undeniable claim that “theoretical terms are introduced explicitly in the context of giving an account of the evidence the theory is meant to accommodate”. This is supposed to support the conclusion that the meaning of theoretical terms cannot transcend all possible evidence: certain facts about the introduction or these terms are supposed to constrain their meaning.

First, consider the question of how exactly theoretical terms serve to “accommodate” (i.e., explain) the evidence. Unless we have some fairly detailed answer to this question, we do not know which constraints, if any, are imposed on the meaning of theoretical terms by the circumstances of their introduction or use. Now one natural answer to this question is that theoretical terms play an explanatory role insofar as they serve to posit unobserved, and perhaps even unobservable, entities. Of course, strictly speaking, individual terms explain nothing. But they can contribute to explanations through their roles in sentences that, together with many others, express certain ontological commitments—to the existence of molecules, for example. The suggestion,
then, is that we “accommodate” (i.e., explain) the evidence by positing evidence-transcendent entities; and there is no obvious reason why the totality of possible evidence might not be best accommodated in this way.

As an illustration of this point, consider Lepore and Ludwig’s claim that, on the present theory of meaning, a theory accounts for the relevant evidence if and only if it makes true predictions about “facts in the domain of evidence”. Thus, a theory of an agent’s degrees of belief and desire accommodates the relevant evidence, i.e., the patterns of that agent’s choice behavior, if it results in true predictions about his choices. The theory might result in predications about how he’ll vote, which route he’ll take to work, and so on. The truth, and hence the meaning, of sentences of the form “X believes that $p$ to degree $n$” is “wholly determined” by the relevant “facts in the domain of evidence” in that the meaning of such claims does not transcend those predictions. The content of such sentences is simply that some agent will behave in certain ways: that he will vote for the Liberals, that he’ll take the Gardiner Expressway, etc. By extension, the meaning of the theoretical terms that figure in these sentences could not transcend whatever contribution they make to the relevant predictions.

It is not obvious that the notion of explanation can be understood simply in terms of successful prediction, but the more important point is that the argument outlined just now is invalid. The premises seem to be these:

1. Theoretical terms are introduced solely so as to account for the “facts in the domain of evidence.”
2. The relevant facts are accounted for by a theory if and only if the theory makes successful predictions about those facts.

But the argument also rests on a tacit assumption:

(A) The way in which a term is introduced (or used) constrains its meaning.

This assumption is somewhat obscure, but it doesn’t much matter. Let us simply assume for the sake of argument that from (1), (2) and some reasonable interpretation of (A), we can infer

3. The meaning of theoretical terms does not transcend their role in the prediction of “facts in the domain of evidence”.

---

103 Davidson, 225.
Nevertheless, (3) does not imply

(4) Predictions about “facts in the domain of evidence” do not express or presuppose commitments to the existence of entities and facts that are not in the domain of evidence.

On the contrary, one might accept (3), and yet hold that in making predictions about the relevant “domain of evidence”, we believe or tacitly assume that there are evidence-transcendent entities—for example, beliefs and degrees of belief, which are not entities directly observable in the choice behavior of agents.

This is a very natural position to take regarding the case of physics or history, as opposed to the controversial case of decision theoretic accounts of belief and desire. Unless we can introduce additional premises about the nature of explanation or prediction, we cannot infer that there are any substantive constraints on the meaning of theoretical terms. Of course, we could make the argument valid by modifying premise (3) so as to rule out the possibility that prediction involves reference to evidence-transcendent entities, but that would beg the question.

In any case, as I’ve stressed, it seems independently plausible that predictions do often rest on commitments to the existence of evidence-transcendent entities. If we predict that Paul will vote for the Liberals on the basis of what he take to be the relations between his degrees of belief and desire regarding various propositions, it seems quite natural to suppose that theory supporting that prediction commits us to the existence of entities not in the domain of evidence, i.e., Paul’s degrees of belief and desire (and perhaps also the propositions he believes or desires true), as distinct from his behavior. Again, this view is especially natural as concerns predictions about “facts in the domain of evidence” relevant to physics, or history.

Of course, it is also possible that the explanatory role of the sentences in which these terms figure is not to express ontological commitments, but to “accommodate” the evidence in some other way—e.g., in some way that, contrary to what we might unreflectively assume, does not require saying anything about an evidence-transcendent world. Certainly this may be the case as concerns some predictions, theories and apparent ontological commitments. For now, I wish merely to emphasize two conclusions. First, no substantive constraints on the meaning of theoretical terms can be
inferred from the plausible claim that they are “introduced explicitly in the context of giving an account of the evidence the theory is meant to accommodate”. This would be true even if they “accommodate” the evidence by occurring in sentences that express ontological commitments to evidence-transcendent entities, which typically say nothing about any kind of evidence. Second, to establish (4), we would need to show that all apparent ontological commitments to evidence-transcendent entities or facts are unreal.

Finally, aside from the question of how the introduction of theoretical terms might “account for the evidence the theory is meant to accommodate”, there is the separate question of which evidence the theory is meant to accommodate. Here it seems that the only motivation for the theory of meaning that Lepore and Ludwig describe leads to some unappealing conclusions. The motivation, presumably, is that the epistemic situation of those who introduce (or, later, make use of) a theoretical term must somehow constrain its meaning. Theorists do not have direct perceptual or cognitive access to molecules, for example, and so, perhaps, the term “molecule” could not have been introduced to refer to such entities. Instead, it must have been introduced to refer to something to which we do have access, certain “facts in the domain of evidence”.

However, the totality of possible sensory evidence is almost as inaccessible to any particular speaker as are molecules, or any other evidence-transcendent entity or fact. So, for the same reason that we are supposed to doubt that the meaning of the term “molecule” could transcend all possible evidence or experience, we should also doubt that it could transcend the evidence actually available to those who introduced the term. The term could not have been introduced, or used on any particular occasion or occasions, with the aim of accommodating all the (supposedly) relevant evidence—all physical measurements and observations, past, present and future, or all patterns of human choice behavior, etc. Nor could it have been introduced or used with the aim of accommodating any significant range of that body of evidence, since no human being could know what that evidence might be.

Now presumably the evidence that the theory (and hence, the introduction of some theoretical term or terms) is “meant to accommodate” must, on any plausible account, be something more than just the evidence actually, presently available to some particular speaker. In introducing the term “degree of belief”, for example, the decision
theorist surely intends to accommodate facts not only about human behavior that he has up to that moment observed, but also facts about human behavior five miles away, or five minutes later, and also facts about what the behavior he has observed would have been if circumstances had been slightly different. (For instance, if at 2:00 the theorist utters the sentence “Paul assigns a subjective probability of 1 to the proposition that the Liberals will be elected next year”, surely Paul’s behavior up until the time of that utterance does not “wholly determine” the truth or falsity of that utterance. His behavior at 2:10, or one week or one year later, might well be part of what makes the utterance true or false.)

The problem, then, is that it seems that the evidence that a theory is “meant to accommodate” on the occasion of the introduction of a theoretical term must allow that this evidence significantly exceeds the evidence actually available, at that time, to those who introduce the terms. Let us assume that the epistemic situation of speakers imposes some substantive constraints on what they could mean by their words. Still, this concession would seem to leave open the possibility that the meaning of theoretical terms sometimes involves facts that are not in the “domain of evidence” at all. The only motivation for denying this possibility was the thought that speakers could not say something about inaccessible entities or facts. On reflection, however, it seems undeniable that they do regularly say such things about inaccessible evidence, at least, and so it is unclear whether there is any principled distinction to be drawn between that kind of inaccessible entity others.

Perhaps evidence inaccessible to a particular speaker, such as evidence that will only be available long after his death, is less inaccessible, in some sense, than evidence-transcendent entities. It is actually impossible that he have that evidence, but, arguably, it is possibly possible that he have it, in that he could have been born a century later. But this seems a fairly weak basis for the radical claim that, although we can say things the truth of which somehow depends on what evidence will be available in a hundred years, we can never say anything about how the world is independently of all evidence.

There is also some sense, arguably, at least, in which it is possible that we have access to facts that actually transcend all the evidence we could possible acquire. For example, perhaps it is logically possible that human beings have greatly expanded sensory powers, which enable them to directly perceive the behavior of molecules. In
that case, it is only actually and not logically impossible that facts about molecules are not “facts in the domain of evidence”, just as it is only actually impossible that we have future evidence of the same kind as evidence that we are (actually) able to obtain.

So, to summarize, there appear to be a number of good reasons to reject the theory of meaning the Lepore and Ludwig ascribe to Davidson, as well as good reasons to doubt that Davidson actually holds such a view. First, as they present it, at least, the theory seems to rest on one of two equally unattractive assumptions: either (a) that there are no special considerations stemming from the nature of meaning and mental life that would warrant a verificationist attitude concerning “interpretation theory”, but not other kinds of theories; or (b) that the same or similar considerations warrant a verificationist attitude toward all theories.

Second, the premise that theoretical terms are introduced explicitly for the purpose of accommodating certain evidence, even in conjunction with the other premises and assumptions noted earlier, does not establish that there are any substantive constraints on the meaning of theories, sentences or theoretical terms. In order to reach that conclusion, we would need the additional premise that the apparent ontological commitments to evidence-transcendent entities of theories supporting the relevant predictions are never the real commitments of the theory. Unless we are to presuppose the conclusion of the argument—that we never refer to evidence-transcendent entities—it is hard to see why we should doubt that prediction is at least sometimes a matter of postulating such entities for the sake of accommodating the evidence.

Moreover, it is very unlikely that we can find acceptable paraphrases of all the sentences that we hold true that eliminate any appearance of quantification over evidence-transcendent entities. So the additional premise needed for the conclusion that Lepore and Ludwig ascribe to Davidson would require denying Quine’s criterion of ontological commitment. But that criterion seems undeniable. How can we make sense of the suggestion that (i) we hold true a sentence of the form “\(\exists x (F_x)\)”, (ii) we have no idea of how to paraphrase that sentence so as to eliminate the predicate, and yet (iii) we do not really believe that anything is in the extension of that predicate?

Finally, the best justification for the theory is the intuition that the epistemic situation of speakers constrains what they could mean by their words. However, this
intuition implies that the idea of the totality of possible sensory evidence (or experience) is only slightly more plausible or intelligible than the idea of a reality transcending all possible evidence. If we take this intuition seriously, it is unclear how the theory can avoid the absurd conclusion that predictions about the very near future, for instance, play no role in determining the meaning of theoretical terms and maintain some principled distinction between evidential and non-evidential facts or entities. In conclusion, we have not found a good reason to accept Davidson’s premise (2), according to which the notion of fitting the totality of sensory evidence or experience somehow implies or reduces to the concept of truth. In particular, premise (2) is not warranted by the kind of argument that Lepore and Ludwig describe.

VII

Let us pause to review the main points of the preceding discussion. I have focused in this chapter on premise (2) in my reconstruction of Davidson’s argument against accounts of conceptual schemes that appeal to the notion of evidential fit: “The thesis that conceptual schemes ‘fit’ some entity implies that alternative conceptual schemes are (largely) true”. We’ve seen that it is highly unlikely that Davidson himself believes this to be true. More importantly, we’ve seen that the argument from meaning empiricism described by Lepore and Ludwig is unconvincing for a number of reasons.

Perhaps there are other, better arguments for premise (2). However, it seems likely that any argument for premise (2) will be open to at least some of the objections set out here. Premise (2) concerns the relation between the concepts of ideal empirical warrant and truth—that is, the relation between the meaning of the claim that a theory fits all possible evidence and the meaning of the claim that it is true. Presumably, then, an argument for premise (2) will aim to show that the first claim is synonymous with the second, or else that it implies the second. The only way to do this, it would seem, is to argue that the meanings of sentences in a theory do not transcend the evidence for those sentences (or for the theory as a whole), and that their truth conditions are therefore restricted to facts about evidence or experience. This general kind of argument will have to involve premises that are not particularly credible, for the reasons given in the
preceding sections of this chapter. But in any case, we’ll see that there are additional reasons to reject Davidson’s case against accounts of differences in conceptual scheme based in the metaphor of fit. So, ultimately, it does not matter much whether I have overlooked some strong argument that might be given for premise (2).

We’ve also noted that it is unlikely that Davidson himself holds an epistemic theory of truth. Rather, it would seem that he attributes some such theory to his opponents, as part of a dialectical strategy. On this reading, which seems to me the most defensible, premise (2) is a step in a *reductio* of the “general position” under discussion, i.e., knowledge empiricism. This is a legitimate move only if some such conception of truth is in fact a key feature of that general position. However, as I suggested earlier, it is not at all clear that his opponents do (or should, given their assumptions) subscribe to an epistemic theory of truth. In this section, I’ll expand on my objections to this attribution of an epistemic theory of truth to proponents of the present account of conceptual schemes.

Quine, who Davidson takes to be one of the most important proponents of the account of conceptual schemes under discussion, claims that “truth is one thing, warranted belief another”. Although it is possible, of course, that he is mistaken about the implications of his own philosophy, Quine at least *thinks* that the concept of truth is not epistemic. Indeed, as we’ve seen, the under-determination of theory by evidence is one of the central themes of Quine’s philosophy, and appears to be one of the main sources of his view that there are many conceptual schemes. But the under-determination thesis is inconsistent with the position Davidson describes.

The under-determination thesis says that no theory can be directly or uniquely derived from the totality of possible sensory evidence. There is always a logical gap between evidence and theory, and so, no matter how strong the sensory evidence for the theory, it is always logically possible that it is false. Since it isn’t possible that more than one of the many empirically equivalent but mutually exclusive theories is true, there are theories warranted by all possible evidence that are, nevertheless, not true. This line of thought is an explicit denial of Davidson’s claim that “if a theory quantifies over physical objects, numbers, or sets, what it says about these entities is true provided the theory as a

---

104 W.V. Quine, *Pursuit of Truth*, 94.
whole fits the sensory evidence” (VICS, 193). Far from encouraging the association of truth to ideally warranted belief, the under-determination thesis seems to invite skepticism, which of course presupposes a radically non-epistemic conception of truth.

Interestingly, Davidson elsewhere remarks that Quine’s position, “like other forms of verificationism, makes for skepticism. For clearly a person’s sensory stimulations could be just as they are and yet the world be very different”. Davidson is right to maintain that there is no logical connection between “a person’s sensory stimulations” and “the world”. This is why, as Quine believes, and as Davidson here seems to agree, even an ideally warranted theory could, in principle, be false; this is what might be thought to lead to skepticism. The problem with verificationist theories of meaning or truth, by contrast, is not that they lead to skepticism, but that they seem to preclude skepticism at the cost of denying any connection between the contents of our beliefs or theories and the external world that we would normally take (most of) them to be about—as with the radical empiricist or verificationist theory of meaning that Lepore and Ludwig ascribe to Davidson. Thus, it is the realist element in Quine’s system that “makes for skepticism”.

Now, as we saw in Chapter 1, it isn’t clear whether knowledge empiricism and the under-determination thesis are consistent with Quine’s empiricist theory of meaning. Quine often appears to identify the meaning or content of an entire theory with the range of sensory stimulation it would cause proponents of the theory to anticipate. This is the “stimulus meaning” or “empirical content” of the theory; and according to Quine, there is no other kind of meaning. If the meaning of a theory does not exceed its “stimulus meaning”, it is incoherent to suppose that a theory “borne out” by “the totality of possible sensory evidence” is nevertheless false. This line of thought certainly seems to be inconsistent with the under-determination thesis and the empiricist theory of knowledge from which it derives.

Whether or not this tension in Quine’s philosophy can be resolved, however, the important point is that, as Davidson recognizes, Quine’s official position appears to be consistent with the skeptical possibility that a theory is “borne out” by all possible

---

evidence and yet false. In characterizing Quine’s view in this way, Davidson is (rightly) conceding that knowledge empiricism does not require an epistemic theory of truth. So there is no obvious reason why Quine could not resolve this tension in favor of knowledge empiricism, and thus strictly distinguish between truth and warranted belief. More importantly, we have found no good reason to assume that knowledge empiricism implies or even encourages meaning empiricism. Quine’s tendency to conjoin knowledge and meaning empiricism may be purely idiosyncratic, and reflect nothing about the nature of these theories—or, by extension, the present account of conceptual schemes. In other words, there may be an account of conceptual schemes, which may be other than Quine’s, which draws on knowledge empiricism, rejects meaning empiricism, and is thus immune to Davidson’s critique.

Moreover, Davidson seems to acknowledge these points. We noted earlier that, according to Davidson, the theory that “experiences, sense-data, surface irritations, or sensory promptings are the sole subject matter of language” (meaning empiricism) is an “extreme, and implausible, version” (VICS, 193) of the “general position” under discussion (knowledge empiricism). So, again, it is hard to see how he could justify his introduction of premise (2). The assimilation of truth to ideally warranted belief would be acceptable only if this “extreme, and implausible” theory of meaning is essential to any instance of the position under discussion (i.e., only if radical meaning empiricism is essential to knowledge empiricism). Otherwise, Davidson’s argument is directed against a straw man.

For various reasons, then, Davidson’s attack on accounts of conceptual scheme that appeal to the metaphor of fit appears to be confused. The position he describes as his target seems to be clearly at odds with the idea of a difference conceptual scheme. Moreover, his description of that position is misleading and perhaps simply incoherent. Quine’s empiricist philosophy would seem to be a key instance of the position under discussion, but Quine’s views are very different from those that Davidson describes. What Davidson needs, but does not provide, is an argument against accounts of conceptual schemes based in a more moderate form of knowledge empiricism that does not comprise radical meaning empiricism. In the next section, I’ll develop an argument of this kind.
I have suggested that the conceptual relativist can maintain a distinction between schemes and empirical content by appeal simply to the distinction between theory and evidence. But in what sense can sensory experience or stimulation be understood as a species of evidence? Although Davidson does not explore this question explicitly in “On the Very Idea of a Conceptual Scheme”, his rejection of empiricism is based in his contention that “nothing can count as a reason for holding a belief except another belief”. Or, to put the point in linguistic terms, nothing can serve as evidence for the acceptance of sentences or theories (groups of sentences) except the acceptance of other sentences.

Sensory stimulation, experience, sense data, etc., are not propositional entities. Indeed, it is for precisely that reason that the conceptual relativist is able to maintain that they are entities distinct from conceptual schemes, theories, sentences, beliefs, etc. But for that very same reason, as Davidson elsewhere emphasizes, sensory stimulation or experience cannot stand in rational relations to theories, sentences, beliefs and other propositional entities. These non-propositional entities cannot be intelligibly taken to mean or justify or imply anything; they are altogether the wrong kind of thing for that cognitive role. Neither, then, can such things be intelligibly described as a species of evidence. Of course, as Davidson acknowledges, it may be that we can make inferences from the fact that a certain episode of stimulation or a certain experience occurs. That something appears red, say, is evidence that it is red. But that a certain kind of stimulation or experience occurs is not itself an episode of stimulation or an experience; it is a propositional entity, the content of a sentence.

Sensory stimulation or experience is typically the cause of our acceptance of sentences or theories. (The mild form of empiricism that Davidson accepts is simply the thesis that the acceptance of sentences or theories is always caused, ultimately, by sensory events.) But a merely causal relation is not an evidential relation. Hence Davidson’s claim that “to speak of sensory experience rather than the evidence” merely

106 “Coherence Theory”, 310-311.
expresses “a view about the source or nature of evidence”: it expresses the view that the cause of the evidence lies in the interaction of our sense organs with the external world. However, it does not “add a new entity to the universe”. For if the present account of conceptual schemes is supposed to rest on the idea of a relation between theory and evidence, the evidence is not sensory experience, stimulation, etc. Instead, it appears that the evidence, strictly speaking, must be something like a certain class of observational or perceptual beliefs. (Or rather, the evidence is a certain class of sentences reporting those beliefs. It does not much matter for present purposes how exactly the evidence is described; the main point is just that it must be propositional, and so cannot be identified with sensory stimulation or experience.)

So far I have simply rehearsed Davidson’s main objection to knowledge empiricism (or knowledge empiricism as traditionally construed). However, if this objection to empiricism is sound, it immediately follows that the present account of conceptual schemes is untenable, apart from Davidson’s discussion of truth and translation. For if the proposal is that conceptual schemes are theories that “fit” the same totality of evidence, and it turns out that the evidence is (say) a class of observation sentences, putatively divergent schemes are simply distinct groups of sentences implied by a single group of observation sentences. (Or perhaps, by inter-translatable observation sentences.) But then the two theories are inter-translatable to some significant extent.

In fact, the two theories are fully inter-translatable. If a certain class of sentences implies each of two different classes of sentences, all three must belong to some single language in which the implications can be expressed. Here it seems there are two possibilities. Either the two putatively different schemes are two formulations of the same theory, differing only notationally, or they are in cognitive disagreement. (The latter possibility could arise, presumably, because of under-determination: it might be impossible to decide which theory to accept because there are different, equally reasonable ways of accommodating the observations reported by the observation sentences.) Obviously notational variations on the same theory cannot be associated with two different conceptual schemes. But, as we’ve seen, cognitive disagreement across conceptual schemes is impossible.
So, this view of the situation satisfies only the first condition on differences in conceptual scheme, that there be some entity common to divergent schemes. It rules out the second condition, that schemes are associated with languages or theories that are not inter-translatable. Essentially, the problem is that, on this view, the evidence in relation to which schemes are supposed to differ turns out to be a fragment of some particular language (and hence conceptual scheme). But then if the languages or theories associated with putatively different schemes stand in rational relations to the evidence, so understood, they must be associated with the same scheme as the first language. The evidence, it seems, is not “something neutral and common that lies outside all schemes”, or some “new entity … against which to test conceptual schemes”. Clearly there is no way to make sense of the idea of a difference in conceptual scheme on these assumptions.

Surprisingly, Davidson does not pursue this objection. As we’ve seen, he attempts, instead, to show that the conceptual relativist’s position rests on an untenable conception of truth as correspondence to sensory evidence or experience. The objection now under consideration, however, is far more compelling. It is independent of any particular conception of truth, and does not impute to the conceptual relativist anything more than the assumption that sensory experience can be characterized as a species of evidence. The upshot is that if some account of conceptual schemes of the kind under discussion is defensible, sensory stimulation or evidence must be characterized as non-propositional entities that nevertheless play the same role as reports of experiences or observations. This does not seem to be a coherent notion.

X

Let us summarize the main conclusions reached in this chapter. The assimilation of truth to ideally (empirically) warranted belief, whether justified on its merits or, instead, as a feature of the present account of conceptual schemes, is a crucial phase in Davidson’s argument. Unless Davidson can show that, somehow or other, his opponents should accept this conceptual assimilation, his contention that the “simple concept of being true” (VICS, 194) implies or presupposes the possibility of translation into “a language we know” (VICS, 194) will not support his conclusion that we cannot make
sense of the idea of a difference in conceptual scheme. Proponents of the scheme idea will be free to reject this initial step in Davidson’s reasoning.

Davidson fails to establish that his opponents should accept this initial step. The assimilation of truth to ideally warranted belief is intrinsically implausible, and can’t be plausibly attributed to his opponents as an essential feature of the general position under discussion. Therefore, we already have sufficient reason to reject Davidson’s attack on the present account of conceptual schemes. However, it is worth considering the remaining steps in his argument, since this may serve to clarify further Davidson’s position and may perhaps suggest other lines of argument for or against the intelligibility of the scheme idea.

In the next chapters, I will explore these other steps in Davidson’s argument against accounts of conceptual schemes based in the metaphor of fit. As we’ll see, they are also open to some important objections. But this is not to say that the positions that Davidson attacks—i.e., knowledge empiricism, on the one hand, and the correspondence theory of truth, on the other—provide an intelligible and defensible basis for talk of differences in conceptual scheme. For reasons other than those that Davidson adduces in “On the Very Idea of a Conceptual Scheme”, these positions would seem to be inconsistent with the notion of a language or theory that is untranslatable, in principle, into our own language. It is not clear why Davidson’s argument takes the form it does, given that the alternative and, to my mind, far more convincing argument developed in the last section is drawn from his own objection to the empiricist position he seeks to attack. So, while I will argue that Davidson’s discussion seems to be confused in some important respects, I will offer other reasons in support of his conclusion that talk of languages or theories “fitting” some entity is of no use to proponents of the scheme idea.
CHAPTER 6
FIT, PART 2: CORRESPONDENCE

In the preceding chapter, I registered a number of misgivings concerning Davidson’s attack on accounts of conceptual schemes based in the metaphor of “fitting”. According to Davidson, the “general position” under discussion is the view “that sensory experience provides all the evidence for the acceptance of sentences” (VICS, 193) (i.e., knowledge empiricism). Conceptual schemes, on this view, are theories (groups of sentences) warranted by “the totality of possible sensory evidence past, present and future” (VICS, 193). Davidson’s strategy against this kind of view hinges on two attempted conceptual reductions:

(i) This concept of “fitting the totality of experience” (VICS, 193) reduces to the concept of correspondence between sentences or theories and something that “makes sentences and theories true” (VICS, 194).

(ii) The concept of correspondence reduces to “the simple concept of being true” (VICS, 193).

Given the premise that the concepts of truth and translation are only mutually intelligible, the proposed account of conceptual schemes qua true theories implies that purportedly alien schemes can be translated into our language.

I’ve argued that the first phase of Davidson’s argument is unsuccessful because he conflates epistemic and metaphysical interpretations of the metaphor of “fit”. To say that a theory fits experience or evidence is just to say that it is warranted by experience or evidence, not that it is true, i.e., that it fits reality, the world or the facts. Perhaps the epistemic interpretation collapses into the metaphysical interpretation in the special case of a theory warranted by all possible sensory evidence, but this is a controversial claim for which Davidson provides no argument. Nor is it clear that all or most proponents of this version of the scheme idea are committed to this position. Quine’s position, for example, is at least prima facie consistent with a firm distinction between the concept of
ideal empirical warrant and truth. So, without some additional argument, Davidson’s reasoning is at best inconclusive.

However, this objection suggests a way in which Davidson’s argument could be repaired. Let us consider again Davidson’s contention that “the notion of fitting the totality of experience … adds nothing intelligible to the simple concept of being true”. This is false as it stands, since sensory experience—a class of sense data, for instance, or Quinean surface irritations—is not a linguistic or conceptual entity, and so it is something with which conceptual schemes (qua theories) could in principle be contrasted. Even if there is no such thing as sensory experience or evidence, so understood, the concept of such experience or evidence “adds” something to the “simple concept of being true” in that, unlike the latter concept, it is the concept of something in relation to which schemes might differ.

Moreover, “the notion of fitting the totality of experience” does “add” something to the concept of truth in that the first is an epistemic notion, whereas that of the second is not. As Davidson says, “to speak of sensory evidence rather than the evidence, or just the facts, expresses a view about the source or nature of evidence” (VICS, 192). (The notion of fitting the totality of experience also “subtracts” something from the notion of truth, of course, since an ideally warranted theory might be false.) If talk of sensory experience expresses a view about the source or nature of evidence—if it involves any claims about evidence at all—it goes beyond talk merely of truth and falsity.

Since Davidson seems to concede at least this second point, his position should not be that the concept of fitting the totality of experience adds nothing to the concept of truth—i.e., that this concept does not allow for any kind of contrast between theories and some other entity that would not be possible simply by appeal to the concept of truth. Instead, his position should be that the concept of fitting the totality of evidence adds nothing to the concept of truth in that it does not allow a certain kind of contrast: a contrast between sentences or theories and something that “makes sentences and theories true” (VICS, 192). Davidson seems to think that if we are to make sense of the present account of the scheme/content distinction, this kind of contrast is needed.

Davidson’s contention, then, should be that the notion of experience or evidence “adds nothing” to the concept of truth in the following sense: the concept of truth does
not provide the needed distinction between *vehicles* and *makers* of truth, and neither does the concept of truth together with some special conception of experience or evidence. This is a more defensible reading of Davidson’s claim that the idea of a fit between sentences and some extra-linguistic entity does not “add a new entity to the universe against which to test conceptual schemes” (VICS, 192).

Despite Davidson’s own characterization of the “general position” under discussion, then, his focus on the idea that “sensory experience provides all the *evidence* for the acceptance of sentences” may be a red herring. Knowledge empiricism figures in Davidson’s discussion only because he associates it with a certain version of the correspondence theory of truth. His position is that “the notion of fitting the facts, or of being true to the facts” (VICS, 192), whether “the facts” are construed in terms of experience or in some other way, reduces to the “simple” concept of truth. If we could make sense of the idea of a correspondence between vehicles of truth and truth makers, the latter class of entities would be “something neutral and common that lies outside all schemes” (VICS, 190). But according to Davidson, the very idea of an entity that “makes” sentences true is either empty or unintelligible. (Presumably this is the point of his reference, in a footnote, to the essay “True to the Facts”, where he argues for this claim.)

This reading is at odds with much of what Davidson says. Most importantly, he describes his target in purely epistemic terms: “talk of fitting” concerns sentences because “it is sentences that predict, … sentences that cope or deal with things, that can be compared or confronted with the evidence”, and “that face the tribunal of experience” (VICS, 193). This is especially puzzling given that, in his earlier outline of the ways in which the scheme idea is typically articulated, Davidson seems to allow two different interpretations of the metaphor of “fit”: conceptual schemes might be said to fit *experience* (sense data, surface irritations, etc.) or *reality* (the world, nature, etc.). As I suggested earlier, it may be that Davidson ignores the second interpretation because it so obviously and directly implies the truth of conceptual schemes *qua* theories. To say that a theory fits reality (or the facts, etc.) may be just to say that it is true, or it may be to offer an explanation of its truth; but in any case, the theory must *be* true. By contrast, it is
less clear that “the notion of fitting the totality of experience” implies truth, and so Davidson argues for this implication.

In any case, whether or not this is what Davidson has in mind in these key passages, we have here the outlines of an argument against certain accounts of conceptual schemes that is immune to the objections considered in the last chapter, and which is therefore worth considering more closely. We can represent this argument roughly as follows:

(1) Alternative conceptual schemes, according to the view we are considering, are theories that “fit” some entity.

(2) To say that a theory “fits” some entity (whether experience, reality, the facts, etc.) is to say that the theory is made true by (or corresponds to) that entity.

(3) To say that a theory is made true by some entity is just to say that it is true.

(4) Ascriptions of truth imply the possibility of translation into our language.

Therefore,

(5) True theories are translatable into our language.

Therefore,

(6) True theories are not alternative conceptual schemes, and the present account of conceptual schemes fails.

(In fact this reasoning was implicit in the outline of Davidson’s argument suggested in the last chapter. However, it is worth recasting the argument in the present form so as to emphasize that, on one reading at least, the argument may be sound even if, as I’ve argued, Davidson’s treatment of knowledge empiricism is confused.) The first phase of the argument thus hinges on the strength of Davidson’s objections to the correspondence theory of truth, even though, surprisingly, these are not explicitly set out anywhere in “On the Very Idea of a Conceptual Scheme”.

It is important to note that Davidson’s case against correspondence theories is not only a subordinate argument in his attack on a certain kind of account of differences in conceptual schemes, which appeals to the notion of a “fit” between theories or languages and something extra-linguistic. In addition, the case against correspondence theories also plays a far more basic and general role in his attack on the very idea of a conceptual scheme, regardless of how exactly the idea is to be understood, and also in his positive
theory of mind and language. In a later paper he writes that his “rejection of facts as entities correspondence to which could explain truth is central to my views of truth and meaning, to my rejection of the scheme-content distinction, representationalism and much more”.  

In the next sections, I’ll evaluate the role of Davidson’s treatment of the notions of fact, correspondence and representation in his attack on the very idea of a conceptual scheme. My conclusion will be that while his case against correspondence theories is fairly strong, it is ultimately irrelevant to the question of whether the idea of a conceptual scheme is intelligible or defensible. Even if one were to accept some form of the correspondence of truth, this would not provide any new basis for the scheme/content distinction. On the contrary, I will argue that any intuitively acceptable notion of correspondence to fact would automatically rule out the possibility that languages or theories corresponding to the same facts are associated with different conceptual schemes.

I

As we’ve seen, Davidson holds that a correspondence theory of truth might be thought to encourage conceptual relativism. If sentences are “made” true (or false) by correspondence to facts, situations or states of affairs, these latter entities may be thought to provide something in relation to which languages or theories (qua conceptual schemes) could differ. This line of thought would not establish that incommensurable languages or theories are possible. But it would at least satisfy one of the conditions Davidson takes to be necessary for talk of differences in conceptual scheme, since it would provide something “neutral and common that lies outside all schemes” (VICS, 190). Thus, if the very idea of correspondence to fact is unintelligible or empty, as Davidson claims, so too is any formulation of the scheme idea that depends on the notion of correspondence, regardless of how one might try to satisfy the second condition (incommensurability).

---

This attack on the idea of a conceptual scheme appeals to the close connection Davidson sees between the notion of correspondence to fact and the notion of conceptual (or mental, or linguistic) representation:

I see no difference between a correspondence view of truth and the idea that utterances (or sentences) ‘represent’, except, perhaps, that if one understood the idea, one could talk of what false as well as true sentences represent. But if there is nothing for true sentences to correspond to, there is nothing for them to represent.¹⁰⁸

Davidson argues that without some notion of conceptual representation, there is no way to make sense of the idea of a difference in conceptual scheme:

Content and scheme … come as a pair; we can let them go together… It is good to be rid of representations, and with them the correspondence theory of truth, since it is thinking there are representations that engenders intimations of relativism.¹⁰⁹

Thus, Davidson identifies “schemes” and representations (although perhaps it would be more accurate to speak of forms of representation). The intelligibility of the scheme idea thus depends on the intelligibility of the correspondence theory of truth. Unless there is something to which true sentences or theories correspond, there is nothing that they represent. Since the very idea of representation is unintelligible, so too is the idea of different, let alone incommensurably different, representations.

The most enthusiastic proponent of this line of thought is Richard Rorty:

If one gives up thinking that there are representations, then one will have little interest in the relation between mind and the world or language and the world. So one will lack interest in either the old disputes between realists and idealists or the contemporary quarrels within analytic philosophy about “realism” and “anti-realism”. For the latter quarrels presuppose that bits of the world “make sentences true”, and that these sentences in turn represent those bits.¹¹⁰

Davidson’s own “anti-representationalism” is more conservative than Rorty’s, but what we should note here is that, like Rorty, Davidson holds that many of the central problems

---

of modern philosophy stem from a representational picture that places “[epistemic] intermediaries between our beliefs and their objects in the world”. If there are any such intermediaries, it is always possible that these intermediaries “could be just as they are and the world be very different”.112

To abandon the notion of representation is to dissolve or at least radically transform many of the central problems of philosophy, including the problem of conceptual relativism. Seen from this perspective, the correspondence theory of truth is virtually identical with the scheme/content distinction. The dualism of scheme and content, according to Davidson, is a dualism of the subjective and the objective, and it seems that, in his view, the dualism ultimately comes down to a distinction between representations and what they represent.

In passing, it is worth noting that, from this perspective, Davidson’s characterization of the “dualism of scheme and content” as a “dogma of empiricism” (VICS, 189) is quite puzzling. Davidson initially characterizes this dualism as a dualism of conceptual scheme and empirical content, as we’ve seen. But however exactly we understand the notion of empirical content, this is not the kind of thing that philosophers would normally take our perceptions, beliefs or sentences to represent. The empirical content of a theory, for example, might be identified with the set of (possible) observations that the theory predicts—or, in Quine’s system, the set of (possible) surface irritations that the theory would lead us to anticipate. And yet it seems odd to say that the theory “represents” those observations or sensory experiences.

On the contrary, if the theory represents anything at all, we would surely want to say that it represents the circumstances in the external world that cause those observations or experiences. So, while empiricism may presuppose representationalism, representationalism would seem to be a far more general and basic thesis. (Descartes, for instance, should presumably count as a representationalist, given Davidson’s account of that position as the view that the identities of beliefs and perceptions do not depend on the environment in which they arise; but he would not count, of course, as an empiricist.

112 “Myth”, 43.
And while Quine is clearly an empiricist, it is doubtful whether he can be fairly described as a representationalist.

So, to summarize, Davidson’s rejection of the correspondence theory of truth has both a local and a global role in his attack on the scheme idea. It provides a subordinate argument for his claim that the idea of a relation of “fit” between theories and some other entity reduces to the indefinable notion of truth, which leads him to conclude that accounts of conceptual schemes based in the metaphor of “fit” ultimately imply the possibility of translation across schemes. But it is also the basis of a much more ambitious and interesting attack on the very idea of a conceptual scheme.

This more general argument can be expressed as follows:

(1) The very idea of a conceptual scheme is intelligible only if the idea of a difference in schemes, i.e., conceptual relativism, is intelligible.

(2) Conceptual relativism is intelligible only if the idea of conceptual (mental or linguistic) representation is intelligible.

(3) The idea of conceptual representation is intelligible only if the correspondence theory of truth is intelligible.

(4) The correspondence theory of truth is intelligible only if there we can identify entities correspondence to which could explain the truth of sentences (or utterances, beliefs, propositions or other vehicles of truth).

(5) We can’t identify any such entities.

Therefore,

(6) The very idea of a conceptual scheme is unintelligible.

II

In this section, I’ll discuss the argument that Davidson takes to be a decisive proof of the impossibility of identifying truth-makers. I’ll then turn to some more general questions about the notion of correspondence to fact, and the relation of this notion to the idea of a difference in conceptual scheme. As I’ve indicated, my suggestion will be that the question of whether the notion of correspondence to fact is tenable or intelligible is not ultimately relevant to the question of whether there could be differences in conceptual
scheme. If anything, a correspondence theory of truth would, if defensible, provide strong reasons to reject the idea of a difference in scheme.

Davidson writes that his reasons for rejecting talk of facts or other truth-makers “are essentially semantical and derive from Frege’s view (and Church’s and G"odel’s arguments) that if sentences correspond to anything, all true sentence and beliefs correspond to the same thing".\textsuperscript{113} The kind of argument to which Davidson alludes here is often called the “slingshot”, and seems to have been inspired by Frege’s view that sentences refer to their truth values. Alonzo Church set out an argument that purports to vindicate Frege’s position.\textsuperscript{114} Davidson deploys a version of Church’s argument in “True to the Facts”.\textsuperscript{115}

Davidson asks under what circumstances a statement of the form

(S) The statement that $p$ corresponds to the fact that $q$

would be true. Presumably a statement of that form is true if both “$p$” and “$q$” are replaced by the same sentence. For instance, a sentence such as

“The statement that Naples is farther north than Red Bluff corresponds to the fact that Naples is farther north than Red Bluff.”

must be true, presumably, if truth is a matter of correspondence to fact. Are there any more interesting or informative substitutions for “$p$” and “$q$” that would preserve the truth of such statements? Davidson claims that since Naples satisfies the description “the largest city within thirty miles of Ischia”, the statement that Naples is farther north than Red Bluff corresponds to the fact that Red Bluff is further south than the largest city within thirty miles of Ischia. But Naples also satisfies the description “the largest city within thirty miles of Ischia, and such that London is in England”. So, “we begin to suspect that if a statement corresponds to one fact, it corresponds to all”.\textsuperscript{116}

According to Davidson, this suspicion can be shown to be correct by appealing to two principles implicit in the foregoing examples:

\begin{itemize}
    \item Alonzo Church, “Review of Carnap’s \textit{Introduction to Semantics}”, \textit{Philosophical Review} 52 (1943): 298-304.
    \item Donald Davidson, “True to the Facts”, in \textit{Inquiries into Truth and Interpretation}.
    \item Ibid., 42.
\end{itemize}
(P1) Statements of the form (S) are true if the sentences replacing ‘p’ and ‘q’ are logically equivalent.
(P2) Statements of the form (S) are true if the sentences replacing “p” and “q” differ only in that a singular term has been replaced by a co-extensive singular term.

Given these principles, we can reason as follows. Let ‘s’ and ‘t’ stand for two arbitrary true sentences. Then the statement that s corresponds to:

(1) the fact that s
(2) the fact that (the x such that x = Diogenes and s) = (the x such that x = Diogenes)
(3) the fact that (the x such that x = Diogenes and t) = (the x such that x = Diogenes)
(4) the fact that t

By a series of purportedly reference preserving substitutions, the argument thus appears to show that all true sentences are co-referential (if they refer at all). The first substitution consists in replacing the second occurrence of “s” with the logically equivalent “(the x such that x = Diogenes and s) = (the x such that x = Diogenes)”: there is no possible situation in which one of these sentences is true and the other is false. The second substitution consists in replacing the singular term “(the x such that x = Diogenes and s)” by the co-referential singular term “(the x such that x = Diogenes and t)”.

Therefore, if any sentence of the form “The statement that s corresponds to the fact that s” is true, so are the sentences resulting from substitutions within that sentence on the pattern of substitutions (2), (3) and (4), and all true sentences have the same referent. For example, if ‘s’ stands for “Plato is a philosopher” and “t” stands for “The earth orbits the sun”, the argument shows that the statement that Plato is a philosopher corresponds to the fact that the earth orbits the sun. Davidson points out that there is no way to individuate facts except by appeal to the sentences or statements to which they are supposed to correspond: facts are theoretical entities posited only for the purpose of somehow explaining or elucidating the notion of truth. But the present conclusion is that
the sentence to which any one fact corresponds is also a sentence to which every other fact corresponds.

So, if the slingshot is sound, there is no way to individuate facts, and all facts collapse into one. Davidson concludes that “descriptions like ‘the fact that there are stupas in Nepal’, if they describe at all, describe the same thing: The Great Fact”:

No point remains in distinguishing among various names of the Great Fact when written after ‘corresponds to’; we may as well settle for the single phrase ‘corresponds to the Great Fact’. This unalterable predicate carries with it a redundant whiff of ontology, but beyond this there is apparently no telling it apart from ‘is true’. 117

The slingshot thus trivializes the notion of correspondence to fact, if that notion was supposed to provide some kind of account of the nature of truth (i.e., of what it is for a sentence or statement to be true). For what could it be that makes all true sentences true, regardless of their subject matter? Seemingly, the only answer is something like “reality” or “the world” or “how things are”.

Of course, there is nothing wrong with saying that a true sentence corresponds to reality (or “the Great Fact”), that it is made true by the world, or by how things are. But this would seem to be at most a way of saying that the sentence is true. This kind of talk does not explain what it is for a sentence to be true. (Although, as we’ll see later, there is perhaps a somewhat less trivial sense in which truth could be construed as a kind of correspondence to reality, which is nonetheless independent of the notion of correspondence to fact.) So, while this reasoning is not made explicit in Davidson’s discussion of the idea that schemes fit some entity, it is the real basis of his assertion that the notion of correspondence “adds nothing to the simple concept of being true” (VICS, 193).

I’ll now consider some objections to this reasoning, mainly for the purpose of clarifying what is at stake, and why Davidson takes the slingshot to have important consequences as regards the idea of a conceptual scheme. As we’ll see, the slingshot may not be as powerful as Davidson has sometimes claimed, but it does seem to generate a serious dilemma for the notion of correspondence to fact. However, as I’ve

117 “True to the Facts”, 42.
emphasized, my main conclusion in this chapter will be that, whatever the force of the slingshot argument, a correspondence theory of truth would not provide a basis for the idea of a difference conceptual scheme. On the contrary, if one were to accept a correspondence theory it would be incoherent to hold that theories or languages corresponding to the same set of facts might be incommensurable.

III

One objection to the version of the slingshot set out in the preceding section is that the transition from (2) to (3) is not reference preserving, or, alternatively, that it is reference preserving only on a question begging assumption.\[118\] Consider the logically equivalent sentences

(A) Plato was a philosopher

and

(B) the \( x \) such that \( (x = Diogenes) \) = the \( x \) such that \( (x = Diogenes & Plato was a philosopher) \).

Sentences (A) and (B) differ in subject matter: they are not “about” the same things. According to this objection, this difference in subject matter may result in a difference in reference. For the second sentence, but not the first, says something about Diogenes. So one might suppose that Diogenes is some part of the fact to which the second sentence corresponds—the fact that Diogenes is Diogenes and Plato is a philosopher, let us say—although no part of that to which the first corresponds. Hence it may seem natural to suppose that, if sentences are referring expressions, logically equivalent sentences may refer or correspond to different entities.

In other words, to simply assume that logically equivalent sentences have the same truth-maker, as the slingshot requires, is just to assume that subject matter is irrelevant to the reference of sentences, and this is effectively to assume the conclusion of the slingshot argument. Of course, this assumption (i.e., (P1), above) is weaker than the argument’s conclusion: it is not equivalent to the claim that the reference of sentences is

\[118\] For an influential development of this objection, see Jon Barwise and John Perry, *Situations and Attitudes* (Cambridge: MIT Press), 1983.
determined simply by their truth values. So the argument does not beg the question, strictly speaking, but it comes very close to doing so, since proponents of the correspondence theory typically assume that subject matter is relevant to the reference of sentences. (By analogy, suppose there is an argument against the existence of natural kinds that begins from the assumption that cats and water molecules belong to the same natural kind. The argument would not beg the question, strictly speaking: it does not depend on the assumption that there are no natural kinds. But it depends on an assumption that is acceptable only to philosophers who do not take seriously the intuitions that motivate talk of natural kinds.)

This objection can be turned aside. For if sentences (A) and (B) are indeed logically equivalent, they must be true under exactly the same circumstances; they have the same truth condition. Sentences with identical truth conditions may differ as concerns other semantic properties—for example, one could perhaps believe (A) to be true but not (B), for some reason, in which case the two sentences presumably differ somehow in meaning. But as concerns the notion of correspondence to fact, the relevant semantic properties are truth conditional. Facts, if such there be, must surely be individuated by their role in establishing the conditions under which the sentences to which they correspond are true. That is, the fact that \( p \) is presumably related to the statement that \( p \) in the following way: the statement is true if and only if the fact exists. A sentence is true, on this kind of view, on the condition that a certain fact exists. So the truth condition of (A) is this: there exists a certain fact \( F \). Given the assumption that (A) and (B) are logically equivalent, we can conclude, first, that their truth conditions are identical. But then the truth condition of (B) must also consist in the existence of \( F \).

Moreover, there are powerful versions of the slingshot that do not require (P1), but merely the weaker assumption that the substitution of some logically equivalent sentences preserves reference. Consider logically equivalent sentences related in the following way:

(i) \( F (\text{the } x \text{ such that } (x = a \& Gb)) \)

and

(ii) \( G (\text{the } y \text{ such that } (y = b \& Fa)) \).

For example, consider these sentences:
(C) (The thing such that it is Plato and the earth orbits the sun) is a philosopher and
(D) (The thing such that it is the earth and Plato is a philosopher) orbits the sun.

These would seem to be just different ways of saying that Plato is a philosopher and the earth orbits the sun. Unlike sentences (A) and (B), sentences (1) and (2) do not seem to be about different objects. Presumably both must correspond to the same fact or facts, if they correspond to anything. That is, we can grant for the sake of argument that sentences (C) and (D) are both about something other than sentence

(E) (The thing such that it is Plato and Diogenes = Diogenes) is a philosopher.

Let us say that (E), unlike (C) or (D), is about Diogenes. So perhaps (E) refers to something other than (C) or (D). Be that as it may, there is no similar basis for denying that (C) and (D) are co-referential. On the contrary, they both seem to be “about” the same thing or things.

But if we allow even this weaker assumption that certain logically equivalent sentences, related in the manner of (C) and (D), are co-referential, we can construct an argument of the following form. Let “Fa” and “Gb” stand for any two arbitrary true sentences. Presumably “Fa” refers or corresponds to

(1) the fact that Fa.

Substituting for the name “a” a co-referential singular term, we can infer that “Fa” also corresponds to

(2) the fact that F (the x such that (x = a & Gb)).

Substituting for the sentence “F (the x such that (x = a & Gb))” a logically equivalent sentence that is about the very same objects, we can infer that “Fa” also corresponds to

(3) the fact G (the y such that (y = b & Fa)).

Finally, substituting for the singular term “the y such that (y = b & Fa)” the co-referential name “b”, we can infer that “Fa” corresponds to

(4) the fact that Gb.\textsuperscript{119}

The move from (2) to (3) requires only the special kind of logical equivalence illustrated above, which does not seem to violate any intuitions regarding what it is that a sentence is “about”.

\textsuperscript{119} This version of the slingshot was suggested to me by Bernard Katz.
So even on the unmotivated assumption that logically equivalent sentences are not always co-referential (if they refer at all), there are still powerful reasons to hold that all true sentences are co-referential. (A parallel argument can be constructed to show that all false sentences are co-referential, but this isn’t important for our purposes.) It seems to me that, in light of these considerations, the onus is on the correspondence theorist to explain why the special kind of logical equivalence that figures in this version of the slingshot is not reference preserving. For it is not easy to imagine why two sentences related in the manner of (i) and (ii), above, should be taken to differ in reference. Again, sentences related in this way would seem to be merely different ways of expressing exactly the same thought or proposition; so one might conclude that they must correspond to the same fact or facts, if they correspond to any. But in any case, the objection considered earlier, which rests on an appeal to an intuitive conception of what it is that sentences are “about”, is not applicable to this version of the slingshot.

Another objection appeals to Russell’s influential theory of descriptions, according to which definite descriptions are not referring terms, but eliminable syncategorogrammatic terms. Using the apparatus of quantificational logic, one can eliminate definite descriptions from the sentences in which they occur; the result of this analysis will be a certain kind of existential generalization. On this kind of view, the slingshot depends on a false presupposition concerning the semantics of descriptions. (P2) says that co-referring singular terms can be substituted without a change in the reference of the sentence within which the substitution occurs. This may be true, but since definite descriptions do not refer to anything, the principle does not apply to the substitution of one definite description for another—as in, for example, the substitution of 

\[ \text{"the } x \text{ such that } (x = \text{Diogenes}) \text{" to "the } x \text{ such that } (x = \text{Diogenes & s}) \text{".} \]

Suppose, then, that we give Russellian interpretations of the sentences “Plato is a philosopher” and “(the \( x \) such that \( x = \text{Plato & the earth orbits the sun} \)) is a philosopher”. The first sentence is to be symbolized by

\[ (1) \ Fa, \]

while the second sentence is to be symbolized by

\[ (2) \ \exists x \ ((Hx & s) & Fx & (\forall y (Hy \rightarrow y = x))). \]
(Where “H” stands for the predicate “is identical with Plato”, “F” stands the predicate “is a philosopher”, and “s” stands for the sentence “The earth orbits the sun.”) Sentence (2) contains no singular terms; therefore, it does not, in particular, contain a singular term that refers to the same thing as “a”. Thus, if (2) reveals the true logical form of the sentence “(the x such that (x = Plato & the earth orbits the sun)) is a philosopher”, the transition from (1) to (2) is not reference preserving, and the slingshot is invalid.

I find this objection unconvincing. To be sure, if we accept Russell’s theory of descriptions, the transition from (1) to (2) can’t be justified by appeal to a principle of substitution of co-refering terms within the context “F_”. Nevertheless, if truth is correspondence to fact, true existential quantifications are just as much in need of truth makers as any other kind of true sentence. An existential quantification is true if and only if there is some object that satisfies the conditions it expresses. In the present case, those conditions are as follows: that there is exactly one thing that is both H (i.e., identical with Plato) and F (i.e., a philosopher). So presumably sentence (2) is made true, if it is true, by those conditions being satisfied, or by the fact that they are satisfied. But this would seem to be simply the fact that Plato is a philosopher.

So Russell’s theory of descriptions does not directly or obviously rule out the kind of substitution needed for the slingshot. Rather, if the theory is true, this might only mean that the semantic situation is somewhat more complex than it appeared, without implying any difference in reference between the sentences in question. To justify the step from sentence (1) to sentence (2) we need only assume that there is some semantic relation between true sentences containing descriptions and the facts that make them true; this need not be the very same semantic relation that holds between sentences containing referring expressions and their truth makers. Perhaps sentences (1) and (2) correspond to their truth makers in different ways, i.e., by picking out an object, on the one hand, and by stating certain conditions satisfied uniquely by that same object, on the other hand. But this kind of difference does not imply that the two sentences correspond to different facts. On the contrary, it would be highly counter-intuitive to suppose that there is one fact that makes true sentence (1), and some other fact—or none at all, perhaps?—that makes true sentence (2).
To be sure, sentences related in the way of (1) and (2) may differ in meaning. There are possible worlds where (1) is true and (2) is false. For example, at certain possible worlds the earth does not orbit the sun, but some other star, but Plato is a philosopher. So, at those worlds, nothing satisfies all the conditions expressed by (2), because nothing is such that the earth orbits the sun. But it remains the case that in the actual world there is a single object that makes (1) true, by being a philosopher, and makes (2) true, by satisfying the condition that there be just one thing that is identical with Plato and is a philosopher. Again, these would seem to be the very same fact. This seems to me an important point, which deserves emphasis. The intuitive notion of correspondence to fact concerns the relation between vehicles of truth and something that makes them true—this latter entity must, presumably, be something actual, and not some merely possible state of things. Hence it would seem that we can safely ignore any differences in meaning that make no difference as concerns the actual world, such as the difference in meaning between sentences (1) and (2). Russell’s theory of descriptions does not provide a particularly compelling objection to the substitution principles that underwrite the slingshot.\textsuperscript{120}

Greg Restall makes a different objection,\textsuperscript{121} directed specifically against Stephen Neale’s reconstruction of Gödel’s slingshot.\textsuperscript{122} The objection is worth considering because Davidson has endorsed Neale’s version of the argument.\textsuperscript{123} According to Neale, the slingshot does not rule out all theories of facts, but places serious constraints on such theories, in that it

\textsuperscript{120} In an influential discussion of Russell’s views, Kurt Gödel hints at a proof of the Fregean view that sentences alike in truth value are co-referential, and notes that Russell’s theory of descriptions would block this line of argument. However, he also expresses some doubt as to the ultimate strength of this response to a slingshot-type argument. It is possible that Gödel had in mind the line of argument sketched here. Kurt Gödel, “Russell's mathematical logic”, The Philosophy of Bertrand Russell, ed., P. Schilpp (Illinois: Open Court, 1944), 125-53.
\textsuperscript{122} Stephen Neale, Facing Facts (Oxford: Oxford University Press, 2001), 183-184. Neale’s slingshot is far more abstract and general than the version we have been considering, but the details of Neale’s version of the argument are not important in the present context.
\textsuperscript{123} Donald Davidson, Truth and Predication (Cambridge: Harvard University Press, 2005), 129.
forces philosophers to say something about the semantics of definite descriptions as soon as they step outside the realm of extensional logic and as soon as they posit entities to which sentences are meant to correspond.  

Specifically, theories of facts must (a) rule out at least one of the principles that underwrites his version of the slingshot, and (b) do justice to the “semi-ordinary, semi-philosophical idea of what facts are”.  

Restall argues that these constraints are entirely trivial, in that any intuitively acceptable theory of facts will automatically rule out at least one of the principles that underwrites Neale’s version of the slingshot. Restall thinks this can be established by considering a possible worlds model for talk of facts. If facts are sets of possible worlds, the transition from (1) to (2), below, fails to preserve reference:

(1) the fact that Plato = Plato
(2) the fact that Plato = the x such that (x = the most famous student of Socrates).

The fact that Plato = Plato, on this kind of view, is a certain set of worlds, the set of all worlds (or, perhaps, the set of all worlds at which Plato exists). By contrast, the fact that Plato = the most famous student of Socrates is a different set of worlds. (There are some worlds at which Plato exists but never meets Socrates, for example.) So, on this view, the substitution of co-referring definite descriptions is automatically ruled out for contexts of fact identity, regardless of what we say about the semantics of definite descriptions.

This is one way to resist the slingshot, but it seems to me that it violates Neale’s second constraint on theories of facts, according to which such theories should do justice to the “semi-ordinary, semi-philosophical idea of what facts are”. One basic intuition about facts is that they are constituents of reality. Facts are supposed to be particular ways things are, particular aspects or features of reality in virtue of which certain sentences or other vehicles of truth are true and others false. Unless a theory of facts somehow captures this intuition, it will not provide entities that can be plausibly understood as “making” sentences true (or false). But Restall’s theory posits countless “facts” that are not constituents of reality, such as the “fact” that Alcibiades was the most

---

124 Facts, 223.
125 Ibid.
famous student of Plato. There is a set of possible worlds at which Alcibiades was the most famous student of Plato, but that entity is not in any sense a constituent of reality.

Restall’s theory also violates a basic intuition about facts in postulating countless “facts” that do not make anything true. On Restall’s account, there is a “fact” that Alcibiades was Plato’s most famous student even though it is false that Alcibiades was Plato’s most famous student. There is nothing in his proposal that distinguishes truth makers from entities that “make” sentences true-at-a-world. Even if we can make sense of the idea that the “fact” that Alcibiades was the most famous student of Socrates makes it true at certain worlds that Alcibiades was Plato’s most famous student, truth-at-a-world is not truth. But it is certainly part of the “the semi-ordinary, semi-philosophical idea of what facts are” that facts are truth makers, not merely makers of truth-at-a-world (whatever exactly that notion might amount to).

IV

Even if Restall’s theory of facts is intuitively unattractive, however, the basic gist of his objection to Neale deserves consideration. Restall’s theory is meant to show that, regardless of how exactly facts are characterized, any intuitively attractive theory of facts will automatically rule out at least one of the problematic substitution principles that underwrites the slingshot. This may be true even if Restall’s account of facts is itself intuitively unacceptable, and, in this respect I am inclined to agree with Restall. The substitution principles necessary for the slingshot are acceptable only if we can assume from the outset that certain intuitions about facts or truth makers can be ignored.

For example, it seems to me that on any natural or intuitively acceptable “idea of what facts are”, we should assume that the sentences

(1) Plato is a philosopher

and

(2) The earth orbits the sun

correspond to different facts. But if so, neither (1) nor (2) corresponds to the same fact or facts as the conjunctive sentence

(3) Plato is a philosopher and the earth orbits the sun.
For presumably sentence (3), unlike (1) or (2), is made true by whatever makes true “Plato is a philosopher” and also by whatever makes true “The earth orbits the sun”. The intuition to which I am appealing is this: there are at least some conjunctions that correspond to more than one fact. This is, of course, equivalent to the claim that there is more than one fact, i.e., that the conclusion of the slingshot is false. But this is not to beg the question: I am not assuming that this intuition is true, but merely that an argument against the notion of correspondence to fact should not begin from the assumption that this intuition is false. Instead, the argument should begin from assumptions that correspondence theorists would be inclined to accept—or at least, assumptions that they would not be inclined to reject out of hand.

Earlier, we noted that the slingshot does not require the assumption that substitution of logical equivalents is generally reference preserving, but only the more modest assumption that a certain class of logically equivalent sentences are co-referential. For instance, if sentences refer, the following sentences are presumably co-referential:

(4) philosopher (the x such that (x = Plato & the earth orbits the sun))
(5) orbits the sun (the y such that (y = Earth & Plato is a philosopher)).

Sentences (4) and (5) seem to be mere paraphrases of (3), and all three of these sentences seem to say the same thing: that Plato is a philosopher and the earth orbits the sun. Since they seem to say the same thing, express the same thought or proposition, it is natural to suppose that they must also correspond to the same fact or facts. All three are made true by the fact that Plato is a philosopher and the fact that the earth orbits the sun. (It is worth noting, also, that it is just this intuition that makes this second transition in the slingshot seem acceptable.) But obviously this is not true of sentences (1) and (2). So presumably neither (1) nor (2) correspond to the same fact as either (4) or (5), since they would then correspond to the same fact as (3).

However, the principle of the substitution of co-referential terms says that (1) and (4), and (2) and (5), are all co-referential, since (4) comes from (1), and (5) from (2), by substitution of co-referential terms. This is highly counter-intuitive. The truth of (4) depends in part on the truth of “The earth orbits the sun”, which has nothing to do with the truth of (1). Seemingly, the fact that the earth orbits the sun (whatever exactly that might be) is part of what makes (4) true, and no part of what makes (1) true. To assume
that there is no important difference in such cases is just to assume that the intuitive “idea
of what facts are” need not be taken seriously. Moreover, if we allow the additional
assumption, which seems quite natural, that (3), (4) and (5) are co-referential, it follows
that (1) and (2) correspond to the same fact as (3), and that (1) and (2) are therefore co-
referential.

It’s interesting to notice that this principle alone would seem to be sufficient for a
collapsing argument without need for any other principle of substitution. If the principle
is true, any true sentence of the form “Fa” refers to the same thing as any sentence of the
form

\[ F(\iota x (x = a \& S1 \& S2 \& \ldots Sn)) , \]

where S1-Sn comprise any conjunction of true sentences. Thus, if we can move from
“Plato is a philosopher” to “philosopher (\iota x (x = Plato \& the earth orbits the sun))”, the
fact that Plato is a philosopher is identical with the fact that makes the second sentence
true—presumably, the fact that Plato is a philosopher and the earth orbits the sun,
whatever that might be.

Now, if some putative fact F1 is identical with some other putative conjunctive
fact comprising that same fact and some other fact F2, F1 and F2 must be the same fact.
Since any and every true sentence can be put in the scope of the iota operator, the
conjunction of “Plato is a philosopher” and every other true sentence corresponds to the
same fact as “Plato is a philosopher”. So if ‘p’ abbreviates any true sentence, including
the conjunction of all true sentences, the fact that Plato is a philosopher is identical with
the fact that p, and there is at most one fact. (If it is problematic to speak of all true
sentences, we can say instead that ‘p’ abbreviates a conjunction of, say, a billion true
sentences. Obviously this more modest conclusion is enough to destroy any theory of
facts.) Notice also that this argument does not rely on any assumptions about the form of
sentences, but needs only the assumption that there is at least one true sentence that can
be represented in the form “Fa”. So the present argument is more powerful than the one
we have considering: it relies only on one assumption about reference and discards the
controversial assumption that all sentences, or even most, can be recast in a single logical
form.
For these reasons, philosophers sympathetic to the intuitive “idea of what facts are” will reject the principle that the substitution of co-referential terms is reference preserving in contexts such as “The fact that $p = q$” or “The statement that $p$ corresponds to the fact that $q$”. But this suggests a serious dilemma for the notion of correspondence to fact. Suppose that the singular term

(T1) the fact that Plato is a philosopher

refers to one thing, while

(T2) the fact that $(x = \text{Plato} \& \text{the earth orbits the sun})$ is a philosopher,

refers to some other thing—say, a complex fact consisting of the fact that Plato is a philosopher and the fact that the earth orbits the sun. If so, it is hard to maintain that the referent of either singular term is a complex entity having $\text{Plato}$ as a constituent. But unless the referents of such singular terms are conceived of in this way, they are indistinguishable from true propositions.

If Plato is a constituent of the referent of (T1), that is presumably just because some constituent of (T1) refers to Plato—just as, if the expression “John and Mary” refers to both John and Mary, that is just because each name refers to a particular person. Of course, the referring constituent here can only be the name “Plato”. But since “the $x$ such that $(x = \text{Plato} \& \text{the earth orbits the sun})$” also refers to Plato, (T2) should refer to the same complex entity as (T1). (Just as, if John is Mary’s husband, “Mary’s husband and Mary” refers to the same thing or things as “John and Mary”.) So, if we reject the principle that substitution of co-referential terms preserves reference for singular terms that appear to refer to facts, we are under pressure to give up the assumption that facts have as constituents the objects about which true sentences say something.

But if facts do not have these objects as constituents—e.g., if Plato is not a constituent of the fact that Plato is a philosopher—it is mysterious how facts could be constituents of reality, as intuition requires. We can ignore the question of what exactly it means to say that something is a “constituent of reality”. My point is simply that facts are supposed to be constituents of reality in some more robust and concrete sense than propositions or sets of possible worlds. It is precisely because of this that facts, unlike those other entities, are supposed to be suited to the role of truth makers. But on the present view they are no better suited to this role than propositions; indeed, it is unclear
how facts, on this view, could be distinguished from true propositions (or, if there is a difference, the sets of possible worlds to which the actual world belongs).

To summarize, the intuitive “idea of what facts are” is motivated by a prior intuition about truth, which would be hard to deny: a sentence or other vehicle of truth says that reality is a certain way, and it is true if and only if reality is as it says. The sentence “Plato is a philosopher” says that Plato is a philosopher, and so it is true if and only if Plato is a philosopher. Facts are supposed to be particular ways things are. In the most basic cases, where a sentence says something about an ordinary, spatio-temporal object, it is that very object that is said to be a certain way, e.g., to be human, to be wise, or to be a philosopher.

So presumably the fact that makes the sentence true (or false) should somehow contain or involve the object about which something is said. Otherwise, the notion of correspondence to fact will not capture the intuition that truth is some kind of relation between the way things are said to be and the way that reality is. But as we’ve seen, this characterization of facts naturally suggests that singular terms referring to facts are insensitive to the substitution of co-referential constituents.

There is thus a deep tension in the very idea of correspondence to fact, since facts can serve as truth makers only if they are characterized in a way that invites the collapse of all facts into one. If singular terms that appear to refer to facts don’t allow substitution of co-referential constituents, facts are indistinguishable from true propositions; they are then vehicles of truth not truth makers. But if such terms do allow substitution of co-referential constituents, there is at most one fact. In either case, there is no reason left postulating facts: in the first case, because entities indistinguishable from true propositions are vehicles of truth and not makers of truth; in the second case, because a single, undifferentiated truth maker adds nothing to the pre-theoretical notion of truth.

V

---

126 Davidson makes essentially the same point, in passing, in his reply to Neale, although he does not explain the objection in detail.
I now turn to the broader question of how the notion of correspondence to fact is related to the idea of a conceptual scheme. Assuming that the objection to correspondence theories advanced in the last chapter is sound, what are the implications for the idea of a difference in conceptual scheme? Earlier, we saw that Davidson thinks that correspondence theories of truth are closely associated with at least some formulations of the idea of a conceptual scheme. In some writings, he takes the stronger position that the notions of conceptual scheme and representation are identical, and that the intelligibility of the scheme idea therefore requires a correspondence of truth. But it seems to me that even the weaker claim of a close association between the notions of correspondence to fact and conceptual scheme is confused. A correspondence theory of truth would not encourage conceptual relativism, and is certainly not a necessary ingredient in the very idea of a conceptual scheme. On the contrary, as can be easily shown, the idea of correspondence to fact is inconsistent with the idea of a difference in conceptual scheme.

Recall that, on the account of conceptual schemes we have been considering in this chapter, the relation of correspondence to fact is supposed to satisfy the first of Davidson’s conditions for a difference in conceptual scheme: the condition that there be “something neutral and common that lies outside all schemes” (VICS, 190), in relation to which schemes differ. Thus, conceptual schemes are incommensurably different theories that correspond to a single body of facts. But this is flatly incoherent. Facts are individuated by the sentences to which they correspond, as Davidson notes. In particular, they are individuated by the propositional contents of those sentences. Only an entity with propositional content could “make” sentences or theories true. This is why, as we saw in the last section, the idea of fact is inherently unstable, since it requires that facts be like propositions despite having as constituents the objects of our sentences and beliefs.

So if a sentence $s$, belonging to the one theory or language, corresponds to a certain fact $F$, it must the same content as $F$. (E.g., the sentence “Plato is a philosopher” has the same content as the fact that Plato is a philosopher.) If that same fact $F$ corresponds to some other sentence $t$, belonging to a second theory or language, $t$ has the same content as $F$, and therefore $s$ and $t$ have the same content. Any two theories or languages related in this way are inter-translatable in the strongest possible sense.
Therefore, any such account of conceptual schemes would be plainly self-defeating, since
the part of the account that satisfies the first of Davidson’s conditions for a difference in
scheme directly implies the failure of the second condition, that conceptual schemes are
incommensurable. For whatever exactly the relation of incommensurability is supposed
to be, it obviously could not hold between sentences or theories having identical
contents.\(^\text{127}\)

However, a correspondence theory of truth might be consistent with a different
formulation of conceptual relativism. For perhaps one might hold that divergent
conceptual schemes are “made true” by different sets of facts. Of course, if these
different sets of facts correspond to incommensurable sentences, they must themselves be
incommensurable: each set of facts (or the propositional contents associated with each)
would stand to the other in the same relation as those of sentences associated with
different conceptual schemes. So, for that reason, the facts corresponding to a conceptual
scheme cannot be coherently characterized as entities common to that scheme and others,
i.e., the common entity in relation to which they are supposed to differ. Presumably, on
this kind of view, the common entity necessary for a difference in scheme must be
something else (sense data, for example).

So, assuming for the sake of argument that the notions of correspondence to fact
and conceptual scheme are both internally consistent, which is of course highly
questionable, it is possible that the two can be consistently combined. Indeed, on the
assumption that divergent schemes are possible, a correspondence theory of truth would
imply that divergent schemes correspond to incommensurable sets of facts, for the
reasons noted just now. However, two points are worth noting in this regard. First, on
the kind of view we are imagining here, facts would have no explanatory role in an
account of conceptual schemes. Since the facts in question must stand in the same

\(^{127}\) Perhaps two theories might correspond to the same facts in a slightly different way: one of the two
theories, or both, might correspond to the facts indirectly, by consisting of sentences with contents that are
implied by the facts. Thus, for example, perhaps the sentence “This rose is red” is made true not by the fact
that a certain rose is red, but by some other, more complicated group of facts that jointly imply that that
rose is red. As we allowed in the last section, the sentence “\(\exists x (\text{identical with Plato}(x) \& \text{philosopher}(x))\)”
does not have the same content as the sentence “Plato is a philosopher”, but they presumably correspond to
the same fact. These complications do not affect the conclusion that correspondence to the same facts,
whatever the details, implies inter-translatability in the strongest sense.
puzzling relation as the languages to which they correspond, they do not resolve the paradox of conceptual relativism, but merely reiterate that paradox. If we do not yet understand how interpretation could be impossible in principle across two languages applicable to a single entity, neither do we understand the relation of the facts corresponding to those languages.

Second, it may be even more difficult to make sense of the idea of incommensurable sets of facts than of the idea of incommensurable languages. Languages are incommensurable if (a) it is impossible in principle to interpret one by means of the other, or to integrate both into some third language by means of which the differences could be understood and described, and (b) both languages are in some sense applied to, or directed toward, a single entity. Presumably incommensurable sets of facts would be related in a different but analogous way: (a) it is impossible in principle to describe or conceive of both sets of facts as constituents of a single world, even though (b) both sets of facts are constituents of a single world.

Facts are supposed to be constituents of reality, truth makers rather than vehicles. Since the notion of incommensurability concerns interpretation or understanding, it can’t be defined directly for facts, but must be defined instead by appeal to the incommensurability of the languages to which facts supposedly correspond. If certain facts are “incommensurable”, this can only mean that it would be impossible to describe or conceive of those facts as features or constituents of a single world even though that is what they are. If this is impossible in principle, and not just for some particular kind of thinker, it is hard to avoid the conclusion that the two sets of facts simply are not constituents of a single world. If this is right, the very idea of incommensurable facts is prima facie incoherent in a way that the idea of incommensurable languages is not.

VI

For the reasons considered in the previous sections, the notion of correspondence to fact will not provide an intelligible basis for an account of differences in conceptual scheme. If the notion of representation implies or presupposes that sentences correspond
to facts, as Davidson believes, we can set aside this second construal of the metaphorical claim that conceptual schemes are theories that fit some entity. However, there are different ways in which languages, theories or systems of concepts might be considered representations, although Davidson does not always distinguish clearly between these. We should consider whether one of these other notions of representation might provide a more defensible account of differences in conceptual scheme.

Language and thought might be said to represent facts—i.e., truth makers characterized by the propositional contents of the sentences they make true. This is the kind of representation that Davidson takes to require a correspondence theory of truth; perhaps such representation is, as Davidson thinks, essentially identical with correspondence. But there is also a more innocuous sense in which language and thought can be said to represent objects: the thought that snow is white can be called a “representation” of snow, for example. This is just to say that thought is about snow. This latter notion of representation has little to do with any particular theory of truth, and is untouched by the slingshot. Likewise, we can also say, innocuously, that language or thought represent the world or reality, accurately or inaccurately. This would seem to be merely a way of saying that sentences or beliefs are true or false. It is worth considering whether the idea of a difference in conceptual scheme might be based on some such notion of representation that is not vulnerable to the slingshot.

Stephen Neale has argued that there is a lacuna in Davidson’s argument against the intelligibility of the scheme idea in this respect. He writes that Davidson rejects the idea of some entity that makes sentences true because “he sees just two plausible candidates: the world itself or individual facts”. Although the slingshot addresses the latter possibility, “it is still open to pursue the idea that a true sentence fits, or is made true by, the world, without endorsing the (possibly hopeless) idea that it fits, or is made true by, a particular fact”. Neale concludes that “in order to demolish the scheme-content distinction, Davidson needs (i) an argument against facts, and (ii) an argument

128 Facts, 61.
129 Ibid., 62.
showing that interestingly divergent schemes are not forthcoming on the view that true sentences are made true not by facts but by the world”.130

Neale points out that although Davidson explicitly denies, in “On the Very Idea of a Conceptual Scheme”, that the world makes sentences true, this seems to be inconsistent with many of his remarks elsewhere. Certainly Davidson thinks that truth is objective, in the minimal sense that whether or not a statement is true does not normally depend on whether it is believed, but rather on how things are in the extra-mental, extra-linguistic world. In this sense, he is willing to say that “true beliefs and thoughts are true because of the way the world is: they correctly ‘represent’ the world”.131 Not only does Davidson fail to argue against this latter notion of representation, then; his own position would seem to allow, or perhaps even require it.

Neale gives the following example of the relation between sentences and the world that he thinks might legitimate some notion of representation and “making true” that is invulnerable to the slingshot:

If the sentence “Smith is in London” is true, it is true because of “how the world is arranged”: one of the entities in the world, Smith, is in London. Indeed, this much is given by a T-theory… So the world makes the sentence true in at least this sense: if the world had been arranged differently—i.e., if the things in the world had been arranged differently (for a world to be arranged, the things in it must be arranged)—if Smith were, say, in Paris, the sentence “Smith is in London” would not be true. To deny this would be to drain all content from the concept of objective truth.132

Davidson agrees that there is a sense in which we can say that true sentences represent or even “correspond to” the world, or that they are made true by how things are arranged in the world: “As long as this way of talking isn’t thought to explain anything about the concept of truth, it is harmless and may even make those happy who want to be sure that the truth of empirical sentences depends on something more than words and speakers”.133

130 Ibid., 64.
131 “Myth”, 46.
132 Facts, 62.
133 “Reply to Stephen Neale”, cited in Facts, 63.
Neale and Davidson both hold that this is just a “way of talking” that does not “explain anything about the concept of truth”. To say that sentences represent or correspond to the world in this sense—e.g., that “Smith is in London” corresponds to Smith’s being in London—is really just to say that they are true. For instance, Neale writes:

It is no more illuminating to be told that a sentence is true if and only if it corresponds to the world than it is to be told that a sentence is true if and only if it is true, states a truth, says the world is as the world is, or fits the facts.¹³⁴

Davidson agrees:

saying the world makes some sentences true … is exactly as harmless as saying that a sentence is true because it corresponds to the One Great Fact, and just as empty.¹³⁵

According to both Neale and Davidson, then, the claim that a sentence “represents”, “corresponds” to or is “made true” by the world, in the sense that they allow, simply means that it is true. That is why Davidson claims that it is “empty”, and does not “explain anything about the concept of truth”, to say that a sentence is true because it represents the world: this is just to say that it is true because it is true. Likewise, when Neale says that such talk of representation is “no more illuminating” than the tautologies he mentions, he means that it is not at all illuminating: it tells us nothing about the truth of sentences that we would not already know simply by knowing that they are true.

I think this position is either incoherent or rests on an equivocation. Although it would not be illuminating to be told that “Smith is in London” is true if and only if it “states a truth” (for example), that is not what Neale says—or not all that he says, at least. Rather, he says that the sentence is true because “one of the entities in the world, Smith, is in London”: the sentence is made true by its relation to a particular arrangement of objects, a particular condition of the world, and not merely by some relation of correspondence to ”the world” as a whole, or merely because it is true or states a truth. To be sure, one might generalize from Neale’s treatment of this example by saying that whether a sentence is true depends on whether it corresponds to the world. But given his

¹³⁴ *Facts*, 62.
¹³⁵ “Reply”, p.668.
earlier remarks, this more general claim would mean that the truth of sentences typically depends on some particular aspect or condition of the world—not that the truth of every sentence depends vacuously on a totally unspecified, undifferentiated entity, such as how the world is, or simply on what is true.

Let us recall the point of the present discussion as concerns the larger question of whether the notion of representation or correspondence provides a basis for an account of differences in conceptual scheme. Davidson’s contention is that it does not because the claim that a sentence or theory “fits” (i.e., representing or corresponding to) some entity adds nothing to the bare claim that the sentence or theory is true. More precisely, this kind of talk does not “add a new entity to the universe against which to test conceptual schemes” (VICS, 194). A “new” entity would be something over and above whatever entity or entities are said or implied to exist by the bare claim that a sentence or theory is true.

So Davidson’s claim that the particular notion of representation under discussion is “empty” means that, like the traditional notion of correspondence to fact, this notion provides us with no conception of some entity in relation to which schemes might differ over and above whatever is provided by the concept of truth. Therefore, this notion is “harmless” in the sense that it could not provide any new basis for an account of conceptual schemes. At least, this is what Davidson’s position regarding this notion of representation should be taken to mean if it is to be relevant to the question of whether some account of conceptual schemes can be drawn from the notion of representation that Davidson allows.

But then the sanitized notion of representation that Neale and Davidson concede to be intelligible has incompatible properties. For it can’t be that the very same notion of representation provides nothing “against which to test conceptual schemes”, and yet also somehow indicates the dependence of truth on something other than words and speakers by relating true sentences to conditions of the world or arrangements of objects. If the claim that a sentence represents the world, in this sense, means or implies there is something other than words and speakers, on which the truth of sentences typically depends, that is thing is also something in relation to which conceptual schemes might be thought to differ. (At least, some reason is needed for the conclusion that this notion of
representation does not provide a basis for the idea of a difference in scheme other than the claim that it is “empty”.) But if the claim that a sentence represents the world, in this sense, does not mean or imply that there is something of that kind, it cannot reassure us that truth is “objective” in the sense that it depends on something other than words and speakers. The notion of representation under discussion may be empty and thus harmless, or it may be reassuringly objective, but it can’t be both.

Now it may be, instead, that there are really two different notions of representation or correspondence under discussion. On the one hand, there is the notion of correspondence to truth conditions or arrangements of objects that Neale draws from the features of a Tarskian truth theory for a natural language. This notion of correspondence seems, at least, to indicate something substantive about the concept of truth as it applies to sentences. It is this notion of correspondence that indicates the objectivity of truth, its dependence on something more than words and speakers. On the other hand, there is the notion of correspondence to “the world”, which, according to Neale and Davidson, is just the notion of sentence that is true, that states a truth, etc. If these are distinct notions, Neale and Davidson are making an invalid argument. For even if claims about this second kind of representation amount to nothing more than claims about the truth of sentences, and are therefore harmless and empty, it does not follow that claims about the first kind of representation are also harmless and empty.

It seems likely to me that this second interpretation is correct. Neale characterizes a certain kind of representation or correspondence, by appeal to a Tarskian truth theory for English, which he clearly takes to license some claim about truth more substantive and specific than the vacuous claim that sentences are true if and only if they are true, state a truth, say that the world is as it is, etc. But he then conflates the purportedly substantive claim with the vacuous claim in order to conclude that this notion of representation cannot underwrite some account of conceptual schemes other than those that Davidson has discredited. As we’ve seen, Davidson appears to accept this invalid reasoning. In later chapters, I will have more to say about the relation of T-sentences to some substantive notion of representation.

VII
What matters for our purposes is that neither the substantive nor the vacuous reading of Neale’s claim that true sentence “represent” or “correspond” to something will fill this gap in Davidson’s case against conceptual schemes. If the notion of representation that Neale and Davidson concede to be intelligible is construed substantively, Davidson is committed to the existence of something that makes sentences true, and he needs to give a reason for his contention that this is not something that could be common to divergent conceptual schemes. On this interpretation, the reason given begs the question. For it may be that this notion of representation is “empty” in the sense that it is identical with the concept of truth, or perhaps with some feature of that concept, but we have no reason as yet to suppose that that concept does not permit a contrast between conceptual schemes and truth makers. The argument against correspondence theories does not apply here, so some other argument is needed.

On the other hand, if the notion of representation under discussion is construed vacuously, divergent conceptual schemes may be possible. For on the vacuous reading, the claim that a sentence represents the world does not mean or imply that there is something with which sentences—or, by extension, conceptual schemes—can be contrasted. But if so, there is no reason to suppose that truth implies the possibility of translation, and the next phase of Davidson’s argument can be rejected. After all, the connection between truth and translation stems from the assumption that if some sentence is true, there must be some particular condition that obtains, which can then be characterized by means of some other true sentence in another language. If ascriptions of truth do not imply that such conditions obtain, or, more generally, that there is something represented by true sentences, there is no reason to suppose that true sentences can be translated or interpreted by other true sentences. For, on this view, there is no reason to suppose that there is something other than (true) sentences that provides a basis for translation or interpretation.

Perhaps the vacuous reading should be, instead, that there are no particular conditions or arrangements of things that sentences represent. What all true sentences represent is simply “the world” or “reality”, where this is taken to be a totally unspecified and undifferentiated entity. This reading is even worse than the first as concerns the
possibility of divergent schemes, for it allows that there is something true sentences represent, and therefore something in relation to which schemes might be thought to differ, but denies that any determinate character. So we have no reason to think that what is represented imposes any determinate constraints on the concepts that can be applied to it, and if there are no constraints, perhaps some of the concepts applicable to this thing are incommensurable.

Of course, if we knew that the concepts of truth and translation are only mutually intelligible, as Davidson argues, we would have a reason to reject this formulation of the scheme idea. But his argument for that conclusion proceeds by attempting to rule out conceptions of truth that do not imply or presuppose the possibility of translation into our language—a subordinate argument against the idea that truth is a relation between sentences and experience or evidence, and a different subordinate argument against the idea that it is a relation between sentences and facts.

As Neale points out, Davidson does not provide an argument against the proposal that truth is, instead, a relation between sentences and “the world”, however exactly that is to be understood. This would seem to be a distinct notion of truth, which, moreover, does not seem to imply or presuppose the possibility of translating any true sentence into our language. So Davidson does not have a case for the inter-dependence of the concepts of truth and translation unless he has a subordinate argument against the proposal that true sentences are made true by correspondence to an unspecified, undifferentiated world. Obviously the subordinate argument should not depend on the premise that those concepts are inter-dependent.

Davidson must choose between a substantivе and a vacuous interpretation of Neale’s claim that true sentences represent or correspond to the world. On either interpretation, the position that Neale and Davidson defend leaves open the possibility of an account of differences in conceptual scheme other than those to which Davidson’s arguments apply. I think that the first horn of this dilemma is preferable to the second. Davidson should acknowledge that there is some (non-tautological) sense in which true sentences represent, correspond to or fit the world, as Neale claims, and which is not vulnerable to the slingshot. This would then require that he give an argument against the
possibility that this substantive notion of representation allows for the possibility of differences in conceptual scheme.

One reason for choosing the first horn over the second is that, as Neale says, the idea of a relation of some kind between sentences and truth conditions, or arrangements of objects, seems to capture the objective content of the concept of truth. If we take the slingshot to rule out the possibility of correspondence to fact, this horn of the dilemma allows us to say that there is nevertheless a reasonably clear and substantive sense in which truth depends on how things are in the objective world. A conception of truth that does not allow for some such dependence must be wrong, since it is precisely this “objective” quality that distinguishes truth from belief.

Another reason is that this interpretation of Neale’s claim that true sentences represent the world can be easily shown to be inconsistent with the idea of a difference in conceptual scheme. Sentences represent, fit or correspond to conditions or arrangements of objects in the present sense only through the “referential apparatus of language” (VICS, 193). For example, the sentence “Smith is in London” can be said to represent, fit or correspond to Smith’s location in the sense that the singular term “Smith” is true of Smith, and the predicate “x is in London” is true of things that are in London. If Smith is in London, that singular term picks out an object that is in the extension of that predicate, and the sentence represents, fits or corresponds to Smith’s location (and London’s).

The sanitized notion of representation that Neale distinguishes from the notion of correspondence to fact is therefore a kind of conceptual or linguistic organization, and so, for reasons considered in Chapter 3, it any two languages representing the same conditions or arrangements of objects are, in principle, interpretable by means of a single, more comprehensive language. (This more comprehensive language may, of course, be one or both of the two languages in question.) So, contrary to Neale’s initial suggestion, Davidson has already given an argument against accounts of conceptual schemes that appeal to the kind of representation or correspondence that Neale describes. Although Neale and Davidson are confused about the nature and status of the idea that true sentences represent the world, the case against the intelligibility of the scheme idea can accommodate this kind of representation or correspondence.
In this chapter, I’ve defended some conclusions that echo Davidson’s, while rejecting much of the reasoning that he advances for those conclusions. Davidson is right to deny that the idea of correspondence to fact provides a basis for the idea of a difference in conceptual scheme. As we’ve seen, the slingshot generates a dilemma for correspondence theories: either the entities that are supposed to make sentences true are characterized in a way that leads to ontological collapse, or they are characterized in a way that makes them indistinguishable from true propositions. On this point, at least, Davidson’s reasoning seems to be sound. If conceptual schemes are supposed to be theories that fit the same facts, there is at most one conceptual scheme, because there is at most one fact.

However, as we’ve seen, the kind of view that Davidson appears to have in mind in his attack on this version of the scheme idea is obviously self-defeating, regardless of whether there is any non-trivial and intelligible notion of correspondence to fact. A correspondence theory of truth is certainly not a necessary condition for conceptual relativism, as Davidson sometimes claims. At most, such a theory would be a very uncomfortable addition to a theory of conceptual schemes, but it may well be inconsistent with the idea of a difference in conceptual scheme. Facts are individuated by the propositional contents of the sentences to which they correspond, and so incommensurable theories could not correspond to the same set of facts.

So there is no need for the slingshot, because, regardless of whether one accepts the slingshot or not, it is obvious that conceptual schemes cannot be characterized as theories that fit the same facts. It is surprising that Davidson does not make this latter objection himself, since it concerns an obvious and striking instance of the “underlying paradox” of conceptual relativism:

Different points of view make sense, but only if there is a common co-ordinate system on which to plot them; yet the existence of a common system belies the claim of dramatic incomparability. (VICS, 184)

One might wonder about the strength of this metaphor as concerns some formulations of the scheme idea. If the “common co-ordinate system” is simply a collection of objects,
for example, it is not obvious that this common system is sufficient to preclude the kind of “dramatic incomparability” envisioned by the conceptual relativist—although, of course, some kind of comparison will have to be possible. By contrast, if the common system is defined by appeal to a set of meanings shared by each of two sets of sentences, on the one hand, and their truth makers, on the other, this instance of the paradox is openly contradictory. Different points of view, in this instance, obviously could not be anything more than differences of opinion.

Earlier, we saw that Davidson’s case against correspondence theories plays two different roles in his attack on the scheme idea. On the one hand, it is intended to provide a subordinate argument for the conclusion that nothing makes sentences or theories true, which is, in turn, a premise in his argument against a specific account of conceptual schemes based—he thinks—in knowledge empiricism. On the other hand, it also seems to provide a far more general and basic argument against the very idea of a conceptual scheme, however exactly the idea is formulated. Conceptual schemes are supposed to be representations, or forms of representation; according to Davidson, the slingshot proves that there are no such things, because there is nothing to represent.

In light of the preceding discussion, neither of these arguments is particularly convincing. We’ve noted already that the subordinate argument is at best superfluous, since a more straightforward argument is possible against the particular characterization of conceptual schemes under discussion. This argument is clearly better than Davidson’s, since it does not depend on controversial or technical assumptions about semantics, nor on any assumptions that go beyond what is implicit in the very idea of correspondence to fact. So, unlike Davidson’s argument, the argument from the shared meanings of sentences corresponding to the same facts leaves no room at all for any defense of this particular account of conceptual schemes.

The more global argument from the slingshot to the conclusion that there are no representations can also be ignored. The only conclusion that can be validly inferred from the slingshot is that there are no representations of a certain kind: no representations of facts, i.e., truth makers having the same propositional contents as the sentences (or propositions, beliefs, etc.) that they make true. The idea that divergent schemes are representations of some single set of facts is just one formulation of the
scheme idea. Indeed, given that it is so obviously incoherent, it is doubtful whether it is a formulation to which any philosopher has been committed. In any case, the very idea of a conceptual scheme, which is supposed to be the target of this more global argument from the slingshot, cannot be identified with the idea of a representation of facts.

There is a sense in which conceptual schemes can be characterized generally as representations, or forms of representation, but this may mean simply that (a) divergent schemes are associated with different ways of thinking about or conceiving of some common thing, or (b) divergent schemes are associated with different ways of organizing, categorizing or classifying the same things. These other notions of representation are untouched by the slingshot, and so the global argument fails except as directed against a self-refuting position that, arguably, no philosopher has ever defended. (Perhaps I should stress that, as we’ve seen, one might combine a correspondence theory of truth and a theory of conceptual schemes.)

Finally, although there seems to be a notion of representation or correspondence that is intelligible even if the slingshot is sound, as Neale points out, this latter notion does not provide a basis for an account of differences in conceptual scheme. Davidson is right to hold that this kind of representation is innocuous, although, again, his reasoning for this conclusion is confused. The reason why talk of this kind of representation is innocuous is not that it as vacuous as the claim that true sentences are true, but rather that it can only be understood as talk about reference, i.e., conceptual or linguistic organization. To say that true sentences represent, in this sense, or to give an account of the semantic relations in virtue of which they do so, is not, perhaps, an explanation of the truth. But it is, at least, a kind of partial elucidation of the concept of truth as it applies to sentences.
CHAPTER 7
TRUTH AND TRANSLATION

In the last two chapters, I’ve argued that the first phase of Davidson’s argument against characterizations of the scheme idea based in the notion of fitting the totality of evidence is unconvincing, for several reasons. The argument does not establish that the notion of fitting the totality of experience or evidence reduces to the simple concept of truth, or that this reduction is plausible given the account of conceptual schemes under discussion. Moreover, Davidson does not offer any clear explanation of how the position he describes, i.e., some form of knowledge empiricism, could be taken to encourage the view that differences in conceptual scheme are possible. Although the very different notion of fitting the facts, or the world, clearly does involve truth, and perhaps reduces to the concept of truth, Davidson’s treatment of this idea rests on a misunderstanding of its relation to the idea of a conceptual scheme.

So Davidson’s discussion of what he takes to be the second main formulation of the scheme idea is in many key respects confused and misleading. But we’ve seen that there are many good reasons to accept his major thesis that the idea of a conceptual scheme, so understood, is not intelligible. Many of these reasons are independent of Davidson’s arguments against the intelligibility of the scheme idea, and internal to the views he takes to encourage conceptual relativism. This makes it even more likely that these formulations of the scheme idea are unintelligible, as Davidson claims, and not just implausible.

In this chapter, I’ll consider the second phase of this argument, which is intended to show that the concepts of truth and translation are inter-dependent. Since the earlier steps in the argument are deeply flawed, as we’ve seen, the argument as a whole cannot establish its conclusion. However, the argument that Davidson makes in the second phase can be separated from this argument against a particular characterization of conceptual schemes, which appeals to the metaphor of fit. We can, instead, take
Davidson’s target to be simply the view that there could be true (or largely true) theories or conceptions of the world that are not intelligible by means of our language.

This more general argument deserves consideration because an alternative scheme, however exactly it is characterized, must presumably have its truths, although these are incommensurable with ours. Moreover, recall that the arguments we have considered in these chapters are all subordinate arguments for premises in Davidson’s master argument against the intelligibility of the very idea of a conceptual scheme. In one way or another, they are intended to support the premise that there is no criterion of language-hood that does not imply or presuppose the possibility of translation (broadly understood) into our language. So, on this reading, this more general argument can be taken to address the proposal that the concept of truth, independently of any connection with the metaphor of fit, might provide such a criterion, and hence an intelligible basis for an account of conceptual schemes.

I

First, let us briefly review the role of Davidson’s contention that the concepts of truth and translation are only mutually intelligible in his larger argument against his intended target, the idea that conceptual schemes fit experience. Since he takes it to be established by the first phase of this argument that “the notion of fitting the totality of experience … adds nothing intelligible to the simple concept of being true” (VICS, 193-4), Davidson concludes that an alternative conceptual scheme, on the present account, is simply a true but untranslatable theory. He argues that we cannot make sense of this idea, because our understanding of the concept of truth as applied to language depends essentially on the concept of “translation into a language we know” (VICS, 195). To suppose that a theory (i.e., a group of sentences) is true is to suppose that it can, in principle, be translated into a language we know.

Therefore, the idea of an alternative conceptual scheme qua true theory is incoherent: an alternative scheme is, by hypothesis, untranslatable into any language we know (or could know); and yet, since it is true, it is translatable into a language we know. As we noted just now, the last steps in this reasoning, if sound, would also provide a
more general argument against the very idea of a conceptual scheme, independent of Davidson’s confused treatment of the particular account of differences in conceptual scheme under discussion. In this section, we’ll examine this reasoning more closely.

Davidson writes that instances of Convention T are trivially true, “yet the totality of such English sentences uniquely determines the concept of truth for English”:

Tarski generalized this observation and made it a test for theories of truth: according to Tarski’s Convention T, a satisfactory theory of truth for a language L must entail, for every sentence s of L, a theorem of the form ‘s is true if and only if p’ where ‘s’ is replaced by a description of s and ‘p’ by s itself if L is English, and by a translation of s into English if L is not English. (VICS, 194)

What Davidson calls “the concept of truth” for a language L is thus the concept expressed by a meta-linguistic predicate of the form “true in L”. The extension of “true in L” for some language L is determined by a class of T-sentences, cast in that same meta-language, and these are, in effect, translations of the sentences of L into the meta-language. (Even if “‘s’ is replaced … by s itself”, the two sentences belong to different languages—e.g., “meta-English” and “object-English”. So, in this sense, every T-sentence can be considered a kind of translation. Of course, “homophonic” translation is not translation in the normal sense of the word, and the difference is important, as we’ll see later.)

Davidson’s point here seems to be that the concept of truth, as applied to language, is intelligible only in relation to some other language, by means of which the truth conditions of sentences of L can be stated. If the concept of truth as applied to language is the concept of truth in L, a meta-linguistic concept, this seems undeniable, and the relation between the mentioned and used sentences that figure in the T-sentences for L is, again, a kind of translation. Therefore, the concept of truth, as applied to language, if this is just the concept of truth in L, is intelligible only in relation to the concept of translation into another language.

This argument is obviously valid. However, Davidson’s ultimate conclusion is not merely that the concept of truth in L is essentially linked to that of translation into some meta-language, by means of which the truth conditions of sentences in L can be stated. Rather, it is that the concept of truth in L is essentially linked to the concept of
“translation into a language we know”, and this doesn’t follow from anything we have seen so far.

In the passage cited above, Davidson writes that, “according to Tarski’s Convention T, a satisfactory theory of truth for a language L” must be cast in English (or some language translatable into English): it must entail T-sentences in which the sentences of L are paired with English sentences that state the truth conditions of the former class of sentences. (An arbitrary sentence s is to be paired with ‘s itself if L is English”, and, otherwise, with “a translation of s into English”.) Now if by “a satisfactory theory of truth for a language L”, Davidson means a theory that English speakers, or speakers of some language inter-translatable with English, are able to understand, this must be true. Obviously a theory of truth, or anything else, for that matter, would not be a “satisfactory” theory if cast in a language we don’t understand.

But while this is obviously a constraint on what we should count as “satisfactory” theories it is no part of the constraint imposed by Convention T. Convention T merely requires that, for any language L, a theory of truth in L entail theorems of the form “‘s’ is true in L if and only if p”—i.e., that sentences of L be paired with their disquotations or translations into some meta-language or other, in the form of T-sentences. This is not to say that the meta-language must be English or Japanese, or any other particular language. Any such theory of truth is “satisfactory” in the formal sense that it “uniquely determines the concept of truth” in or for L, i.e., the extension of the predicate “true in L” (or, as the case may be, some co-extensive predicate in another language). This has nothing to do with whether the theory is “satisfactory” in the sense of being intelligible to English speakers.

So there is nothing unintelligible or incoherent in the idea of a Tarskian truth theory for some language—say, Martian—that is cast in a meta-language that we could not come to speak or understand. For instance, there might be a truth theory that entails the T-sentence in Jupiterian, or simply Meta-Martian, to the effect that the Martian sentence “?” is true in Martian if and only if ?. (I say “to the effect that” so as not to assume that these meta-languages contain the phrase “is true in Martian if and only if”). Such a theory might conform to Convention T, despite not being intelligible or, in that sense, “satisfactory”, to us, or to speakers of any language inter-translatable with English.
Neil Tennant argues that, in light of these considerations, we can make sense not only the idea of a language for which no English truth theory is possible, but also the idea of a conceptual scheme incommensurable with our own. Tennant agrees that the metalinguage in which a truth theory for (say) Martian “need not be English, or any language translatable into English”. Rather, it might be that a truth theory for Martian would be possible, but “only in languages employing the same conceptual scheme”. According to Tennant, we can speak of a difference in conceptual scheme in such cases because the “range of truth-conditions” for some radically alien language or theory would be “truth-conditions-for-them [i.e., the aliens], truth-conditions expressed in their terms, truth-conditions constituted by their concepts”. The alien concepts, or the truth conditions constituted by those concepts, are “incommensurable” with our own, even though the notion of truth is univocal.

For reasons that are by now familiar, I think that this conclusion is unwarranted. What we’ve found is only that Davidson has not established the radical conclusion of his discussion: the mere existence of a Tarskian truth theory for some language L1 expressible in some language L2 implies only that both languages are associated with a single conceptual scheme. It doesn’t follow from this that either L1 or L2 is associated with our conceptual scheme. That stronger conclusion would follow only on the assumption that there could not be group of inter-translatable languages none of which are translatable (or interpretable) by means of our language, which is, of course, precisely what is supposed to be in question here.

However, we do not have any reason, as yet, to think that there could be some such language or group of languages. To say that a truth theory for some alien language would state “truth-conditions-for-them”, or “truth conditions constituted by their [alien] concepts”, which could not be identified with truth-conditions-for-us, is just to assert the

---

137 Ibid., 81.
138 Ibid., 86.
139 Ibid.
possibility of an alternative scheme. It is not to explain how that kind of relation between
ccepts or languages applied to the same world is possible, or what exactly the relation
is supposed to be. In particular, this is no explanation of how the concept of truth could
serve as a basis for incommensurability.

The explanation that Tennant proposes is that the Martian concepts could be
inaccessible to us because they apply to a “species-specific kind of input”—i.e., an
empirical content available only to creatures with perceptual and cognitive systems
radically unlike ours. This would seem to be just the idea, which we considered and
dismissed earlier, that divergent schemes are simply very different systems of concepts
that organize the same entities. But as we’ve seen, mere conceptual diversity is not
enough for incommensurability. The truths expressed in Martian might concern (for
example) properties detectable only with sense organs radically unlike ours, but those
truths could be seen to be complement those expressible in our language, just as truths
about colors complement truths about sounds or tastes.

III

The main conclusion of the preceding sections has been that Convention T does
not impose on ascriptions of truth the radical constraint that Davidson claims. However,
a different reading of the argument is possible, which does not depend on this mistaken
view of Convention T. Let us reconsider Davidson’s reasoning. Davidson claims that a
(roughly) Tarskian truth theory for English sentences “uniquely determines the concept
of truth for English”. The concept of truth in English (or “for”, or “relative to” English)
is “determined” by such a theory in the sense that, if it is empirically adequate, the theory
specifies, for any arbitrary English sentence, the condition under which the predicate
“true in English” applies to that sentence. Generalizing on this example, all concepts
expressed by predicates of the form “true in L” are determined by some such truth theory,
i.e., a truth theory for the language under discussion.

This is why Davidson takes the concept of truth to be inseparable from that of
translation. Predicates of the form “true in L” are applicable only in some meta-language

140 “Logic and Conceptual Schemes”, 85.
or other. The extensions of such predicates are determined only by the relevant set of T-sentences, which must be cast in some language other than the one under discussion, which can be used to translate the latter. So if we then go on to say that certain sentences are true in some language, this is an intelligible claim only if there is a truth theory containing a predicate of the form “true in L”, applicable to those sentences. The truth theory is then a translation manual for the object-language.

Presumably this is the meaning of Davidson’s somewhat mysterious claim that “Convention T suggests, though it cannot state, an important feature common to all the specialized concepts of truth [i.e., all the instances of truth in L] … by making essential use of the notion of translation” (VICS, 195). What is suggested is a conceptual dependence of ascriptions of truth in a language on the possibility of translation: for any true sentence of the form

(S) s is true in L,

L must be translatable into the language of (S). We do not need to assume that we can, presently, translate into our language a sentence that we take to be true in some other language. The point is rather that if our language can be used to apply the concept of truth in a language to some other, translation into our language must be possible, whether we actually have a truth theory for that language or not.

This conceptual dependence blocks the objection made in the last section, that Convention T does not make any reference to translation into our language, or any particular language. The objection could be expressed by saying that, although it might be self-defeating to say that a particular group of sentences is true in some language L, we can imagine merely that some language or other contains sentences true in that language, but not translatable into our language. This does not violate Convention T, since we can allow that those sentences are translatable into some language. This position seems intelligible. We can acquire the general concept of truth in a language by reference to the “specialized concepts of truth” that can be applied in our language; we can then make use of that general concept, independently of those more specific concepts, to imagine sentences that are true, and so translatable into some language, but not translatable into ours.
However, if there is a general concept of truth in a language that subsumes both the concept of truth in our language and the concept of truth, there could also be a truth theory for both of those languages cast in some third language that translates both. It makes no difference whether we claim to identify the true but purportedly untranslatable sentences. As with the argument against the idea that divergent schemes organize the same entities in incommensurable ways, this argument is not verificationist: it does not concern the criteria for recognizing or identifying languages other than our own. The argument is instead directed, as it should be, at the very idea of a difference in conceptual scheme, and aims to show that the concept of a true theory turns involve a commitment to the possibility of translation. So, as Davidson promised, the claim that all languages are by nature inter-translatable emerges as the conclusion of an argument.

The argument from Convention T can thus be seen as another instance of Davidson’s more general strategy against the intelligibility of the scheme idea. Again, this strategy stems from his observation that

the dominant metaphor of conceptual relativism, that of differing points of view, seems to betray an underlying paradox. Different points of view make sense, but only if there is a common co-ordinate system on which to plot them; yet the existence of a common system belies the claim of dramatic incomparability. (VICS, 184)

In the present context, the “different points of view” are theories, i.e., groups of sentences, while the “common co-ordinate system” is provided by the concept of truth.

To illustrate this point, let us suppose that there is a Martian sentence “?”, which belongs to an alien conceptual scheme. So as not to beg the question, we will not assume that the truth conditions of sentences of the Martian language can be framed in any English truth theory. Instead, we will merely suppose that there is some language, Jupiterian, say, by means of which Martian truth conditions can be characterized. There is, then, some Jupiterian sentence to the effect that

(M) “?” is true in Martian if and only if ?. (Again, I say “to the effect that” so as not to assume that the Jupiterian sentence will consist of just those words.) This much follows from the mere fact that, by hypothesis, “?” is true in Martian, since the concept of truth in Martian must belong to some
language other than Martian, by means of which Martian can be discussed. Of course, we also need to say that some of our own sentences are true in our language. This is a necessary condition for a difference in conceptual scheme, on the present account, since it is the concept of truth that is supposed to provide a criterion of language-hood that does not imply or presuppose the possibility of translation: both our language and the alien language contain true sentences, i.e., sentences true in their respective languages.

So we also hold that there is some language by means of which the truth conditions of our language can be framed, as in the sentence

(E) “Snow is white” is true in English if and only if snow is white.

But now suppose that, as was suggested earlier, both of the predicates “true in Martian” and “true in English” are subsumed by some more general predicate “true in Martian and English”. If so, then given the dependence of any such predicate on a meta-language that translates the sentences to which the predicate applies, there must be a third language, e.g., Jupiterian, that can translate both Martian and English sentences.

But this is just to say that Martian and English can be treated as a single language, with a single truth predicate, “true in Marglish”, applicable to the sentences of both. Thus, if the predicates “true in Martian” and “true in English” are taken to be special cases of a more general predicate, the true sentences of both of the lesser languages will be translated by sentences that are true in some single (meta-) language, and the truths they express will be simply different and complementary. This is inconsistent with the idea that they are associated with different schemes, which requires not merely conceptual diversity, but incommensurability. On the present view, the truths expressed by Martian and English sentences are related in the same ways as truths expressed by English sentences with diverse objects—for example, sentences that express truths about colors versus sentences that express truths about money.

In the case of a difference in conceptual scheme, therefore, the truth predicates applicable to the two languages in question can’t be subsumed under a more general truth predicate. Clearly this is just to say, though, that the original suggestion that both the Martian sentence “?” and the English sentence “Snow is white” are true, or that both are true in some language, involves an equivocation. Perhaps “?” is true (or true in a language) in one sense and “Snow is white” is true (or true in a language) in some other
sense. But on this view, there is no longer any “common co-ordinate system”, provided by a shared concept of truth, on the basis of which English and Martian sentences or theories could, in principle, be compared to “different points of view”.

IV

Davidson’s highly compressed remarks do not clearly show whether the kind of argument described in the last section is in fact what he intends. This is, however, an interpretation of his discussion that does not simply beg the question, in a fairly obvious way, against the suggestion that there could be a true but untranslatable theory. In any case, this is an argument from Convention T to the conclusion that there could not be true but untranslatable theories. On this reading, Davidson’s claim that Convention T requires that truth conditions for sentences in the extension of any predicate of the form “true in L” be stated either in English or in a language translatable into English is not a premise, but rather a conclusion. The premise is merely that Convention T requires some meta-language or other for any sentences in the extension of such a predicate. So although Convention T does not actually say anything about English, when it is applied to the present account of conceptual schemes, which concerns our language and some other, it implies the possibility of translation between our language and the other.

Now one important feature of this line of thought, which is perhaps not obvious, is that it does not rest on an appeal to some general concept of truth independent of Convention T. The claim is not that if one sentence is true-in-L1 and another sentence is true-in-L2, they must both have the property of truth, i.e., truth simpliciter, “in” or “relative to” nothing. Indeed, that intuitively attractive assumption would undermine Davidson’s position, since the conceptual dependence of ascriptions of truth on the possibility of translation into a meta-language is based on Convention T. A more general concept, not defined by Convention T, might not have any such dependence on the concept of translation. Instead, the argument depends merely on an appeal to the concept of truth in a language, divorced from any more general or intuitive notion of truth (i.e., truth simpliciter).
What the argument purports to establish, then, is not, as it may seem, that all predicates of the form “true in L” express or depend on some prior concept of what is simply true, in or relative to nothing. Rather, the view is that all intelligible ascriptions of truth are in fact ascriptions of truth in some language or other; since the very concept of truth in a language is essentially linked to the concept of translation into a different language, the idea of a true but untranslatable sentence or theory is incoherent. The more general concept of truth to which the argument appeals is simply a more general concept of truth in some language, the language consisting of all the sentences associated with two supposedly divergent schemes. The conclusion of the argument is therefore that all truth is relative to a language, in particular to our language, in the sense that the very concept of truth is the concept of truth in some language inter-translatable with ours.

I think that this reasoning is unsound, for reasons that I’ll try to explain in this rest of this chapter. My suggestion will be that the assumption that all ascriptions of truth are ascriptions of truth in some language stems from Davidson’s view of the nature of linguistic meaning and understanding, according to which all understanding is interpretation. In my view, both of these counter-intuitive positions reverse the real conceptual priorities: the concept of interpretation is dependent on that of understanding, in some non-interpretive sense, and the concept of truth in a language is dependent on that of truth, “in” or “relative to” nothing. I’ll argue that Davidson’s view is not merely counter-intuitive, but incoherent. It stems from a misuse of the concepts of meaning, truth, interpretation and understanding.

First I will try to explain the basis of these views in Davidson’s philosophy, and why I think they must be rejected. Then I’ll consider how this bears on the present account of conceptual schemes as true but untranslatable theories. As we’ll see, there is a good argument against this account of conceptual schemes that has much the same structure as the argument that I have attributed to Davidson, but which does not rest on these mistaken assumptions about linguistic meaning and understanding. So, as with Davidson’s attack on the idea that conceptual schemes can be characterized as sentences made true by the same facts, the argument from Convention T is at best superfluous. A more straightforward argument can secure the same conclusion, without appeal to Davidson’s controversial and, in my view, mistaken assumptions.
We can approach the more general view of linguistic meaning and understanding that underwrites the argument from Convention T by considering what would appear to be an obvious and decisive objection to the reading of the argument advanced earlier. For perhaps it seems obvious that the argument, even on this non-verificationist reading, fails to establish anything about Davidson’s ostensible target, which was supposed to be an instance of the idea of a difference in conceptual scheme. The argument under discussion establishes that for any two sentences \( s \) and \( t \), if \( s \) is true in \( L_1 \) and \( t \) is true in \( L_2 \), both \( s \) and \( t \) must be true in \( L_3 \), and hence \( L_1 \) and \( L_2 \) are translatable into \( L_3 \). But then suppose that a truth theory for \( L_1 \) and \( L_2 \), framed in \( L_3 \), entails the T-sentence

\[
\text{(T) “s and } t \text{” is true in } L_1 \text{ and } L_2 \text{ if and only if } s \text{ and } t.
\]

The second occurrence of “\( s \) and \( t \)” is, in a certain sense, a translation of the first, if only for the legalistic reason that the second occurrence belongs to the meta-language while the first belongs to the object-language. There is a language under discussion, and a second language being used to discuss the first. So, mere disquotation would count, in this legalistic sense, as a “translation” of \( s \) and \( t \).

Now this kind of translation is surely not sufficient for an interpretation or understanding of either sentence. Unless one already understands what is meant by the meta-language sentence “\( s \) and \( t \)”, which simply disquotes the object-language sentence, \( (T) \) cannot serve to interpret the object-language sentence. Nevertheless, it can be known \( a \ priori \), without any knowledge of what that either conjunct of that sentence means, that \( (T) \) is true—that it correctly states the truth conditions of the mentioned sentence, and, in that sense, translates or gives the meaning of the mentioned sentence. But this could be true even on the assumption, which we may grant for the sake of argument, that \( s \) and \( t \) are associated with divergent schemes. That is, suppose that merely to understand or hold true either of those sentences would make it impossible, somehow, to understand or hold true the other. Even so, \( (T) \) would be true—but \( (T) \) itself would belong to a language that no one could understand. All of this is consistent with the minimal constraints imposed by Convention T on predicates of the form “true in \( L \)”. 

205
In light of these considerations, it is important to note that, as we’ve seen before, Davidson’s appeal to the notion of translation is misguided, even by his own standards. What is really at issue here, or should be, is not the question of whether a pair of true theories could be mutually untranslatable, in the sense of not being subject to a single Tarskian truth theory. For one thing, it should be clear by now that no two languages could be mutually untranslatable in that sense, as the argument set out just now establishes. But that kind of translatability is not relevant to the question of interpretation or understanding: it is not sufficient for identity of conceptual scheme, since even if we suppose that two theories or languages are incommensurable, it will still be trivially true that there is a Tarskian truth theory applicable to both. Identity of conceptual scheme would be implied only by the possibility of interpretive translation, which is not guaranteed by the existence of such a truth theory.

Moreover, as we saw in Chapter 3, the impossibility of translating (in some stronger sense) one language into another is not sufficient for a difference in conceptual scheme, since it may be that speakers of either language can simply learn the other. Nor is the fact that some speaker is incapable of understanding or interpreting a language sufficient for a difference in scheme. This might be impossible for some reason other than the relation between the speaker’s own language or concepts, on the one hand, and those that he is incapable of understanding, on the other—e.g., stupidity, bad eyesight, psychological resistance. Hence the question here should be whether there could be two theories such that merely to believe or understand one would make it impossible to understand the other. For only in that case would we have traced the failure or impossibility of understanding across languages to a special relation between the languages (i.e., incommensurability).

Why, then, does Davidson not acknowledge this key distinction between interpretive translation, which would be sufficient for identity of conceptual scheme, and mere disquotation? The reason, I will argue, is that his theory of linguistic meaning and understanding does not really allow for this distinction between understanding or interpretation, on the one hand, and translation, or even mere disquotation, on the other.

Davidson holds that the facts about the meaning of a speaker’s language are exhausted by an empirically adequate truth theory for his language. Since that meaning
is what a speaker understands, presumably, in knowing his language, it would seem to follow that linguistic understanding is exhausted by knowledge of such a theory. In fact, because Davidson accepts a version of Quine’s thesis of the indeterminacy of translation, he holds that any particular truth theory of that kind will represent the speaker as meaning or understanding more than he really does. So a more precise formulation would be that, for Davidson, the facts about a speaker’s understanding are exhausted by what is invariant across empirically equivalent theories of truth for his language. This view does not seem to leave any logical space for the question of whether a merely disquotational T-sentence could serve as a genuinely interpretive translation.

Let us consider more closely what, if anything, Davidson takes to be the nature of linguistic meaning, of what it is that a speaker knows in understanding his language. One key feature of Davidson’s position is his influential claim that a (roughly) Tarskian truth theory can provide a compositional theory of meaning for a natural language. Because such a theory would determine the meaning of any arbitrary sentence in the language, on the basis of its constituent expressions, knowledge of the theory would be sufficient for understanding the language: “To know the semantic concept of truth for a language is to know what it is for a sentence—any sentence—to be true, and this amounts, in one good sense we can give to the phrase, to understanding the language”. 141

To devise an empirically adequate truth theory for a language would be to interpret the language, in Davidson’s sense of the term. This is not yet to say that linguistic understanding consists in the very same kind of knowledge that an interpreter might achieve by the construction of such a theory. For it could be, of course, that what the interpreter learns in this way is known or understood by speakers of the language in some other way. This suggestion reflects an intuitive view of meaning and understanding, according to which interpretation or translation is a special, derivative kind of understanding, which depends on a prior understanding of one’s own language, which must then be a different kind of understanding.

But this suggestion ignores the central importance of the interpreter’s perspective in Davidson’s philosophy. With Quine, Davidson holds that language is an essentially social and hence public institution. Consequently, there could not be a form of linguistic understanding that is not public.

141 Donald Davidson, “Truth and Meaning”, in Inquiries into Truth and Interpretation, 24.
meaning that transcends what would be discernable to a theorist trying to construct a 
theory of meaning and belief without any prior knowledge of the target language. 
Davidson thus regards the constraints on the radical interpreter’s project as revealing the 
basic features of linguistic understanding generally. The radical interpreter is not doing 
in one, unusual way what native speakers of the language do in some other way; it is just 
that the difficulties and constraints are more obvious to the interpreter.

The evidence available to a radical interpreter is simply the behavior of speakers, 
on the one hand, together with what he takes to be the objective conditions of the shared 
environment, on the other. This is, of course, just the kind of evidence that a truth theory 
captures. Other kinds of theories, invoking other semantic concepts, could not be based 
on the kind of evidence available to a radical interpreter. For example, a radical 
interpreter has no way of discovering which propositions a speaker intends to express. 
So truth theories would seem to provide not only theories of meaning in the narrow sense 
that they can be used to determine what expressions in a language mean, but also in the 
broad sense that the general character of such theories represents the nature of linguistic 
meaning, and hence of understanding.

Davidson’s position thus depends on two sources. First, he believes that the range 
of empirically adequate truth theories exhaustively characterize the semantic facts. This 
view is motivated by fairly abstract logical and semantic considerations, such as, for 
extample, his rejection of intensional entities on the grounds that they do no explanatory 
work in a compositional theory of meaning. Second, he believes that, given the nature of 
language acquisition and use, there can be no facts about meaning or understanding that 
would transcend knowledge available to a radical interpreter. Taken together, these 
views would seem to imply that all linguistic understanding is essentially a process of 
thruth-theoretic interpretation.

Davidson has quite clearly endorsed this conclusion in many key passages, 
although, as we’ll see, there is a serious tension in his philosophy concerning this issue. 
Thus, in “A Nice Derangement of Epitaphs”, Davidson writes that one person’s ability to 
understand another’s speech is “the ability that permits him to construct a correct, that is, 
convergent, passing [truth] theory for speech transactions with that person”,

\[142\] Donald Davidson, “A Nice Derangement of Epitaphs”, in *Truth, Language and History*, 106.
understanding is “the process of creating new theories to cope with new data”. In this sense, then, “all understanding of the speech of another involves radical interpretation”. At times, Davidson appears to suggest that understanding of one’s own utterances also involves interpretation, or something akin to interpretation: “A competent speaker or interpreter is able to interpret utterances, his own or others” only because there are “systematic relations between the meanings of utterances” [emphasis mine].

In fact, this view would seem to be implied by Davidson’s account of meaning. For if a speaker understands his own words in a way that does not involve truth-theoretic interpretation, it would then be the kind of meaning that one knows in that kind of understanding, whatever it might be, that an interpreter aims to represent with a truth theory. But if that kind of understanding does not itself take the form of a truth theory, it might be that the interpreter’s theory of meaning for a speaker’s language fundamentally misrepresents what the speaker means by his words. How could we ever know that what is captured by the interpreter’s truth theory is the very same meaning that speakers of the language understand in some other way? This line of thought would violate the Quinean requirement that knowledge of meaning cannot transcend what an interpreter (or translator) could discover, and that all understanding must hence be the same in kind as the interpreter’s understanding of an alien language.

On Davidson’s view, then, all understanding of linguistic meaning is a kind of interpretation, which rests essentially on a truth theory, and hence on a kind of translation. What I want to emphasize here is that, on this view, the epistemologically and semantically basic form of interpretation is represented by disquotational truth theories. First, consider the question of what difference there could be, on this view, between what an interpreter takes a speaker to mean or understand by his sentences and what the speaker himself means or understands by those sentences. The difference can only be that the speaker, unlike the interpreter, understands—i.e., “interprets”—his own utterances by means of disquotational T-sentences. This is not to say, of course, that the speaker actually carries out a process of “interpretation” in such cases, but that what he

---

143 Donald Davidson, “Radical Interpretation”, in *Inquiries into Truth and Interpretation*, 125.
144 “Epitaphs”, 93.
knows in knowing what his utterances mean, is exhaustively characterized by a disquotational T-sentence (together with the theory that generates it as a theorem).

For example, the interpreter may understand a speaker’s utterance of “Schnee ist weiss” by means of truth theory that entails the T-sentence

\[(T1) \text{“Schnee ist weiss” is true in German if and only if snow is white.}\]

By contrast, the English speaking interpreter’s understanding of his own sentence “Snow is white”, which he uses to interpret the German sentence, consists in his knowledge that

\[(T2) \text{“Snow is white” is true in English if and only if snow is white.}\]

Davidson holds a speaker cannot say anything, to others or to himself, that would characterize in a more informative way what he means by his own utterance—more precisely, the best he can do is to say something of the form “My utterance of ‘s’ is true if and only if s”.\(^{145}\) Similarly, Davidson writes that if two people speak “what is in every empirically determinable respect the same language” they “can (and presumably do) use the homophonic translation manual in understanding one another”.\(^{146}\) So again, it would seem that to understand one’s own language is to possess a disquotational (or “homophonic”) truth theory.

Now, if a speaker’s understanding of his own utterances is best represented as knowledge of a disquotational T-sentence, or can be so characterized, and if that kind of understanding is interpretive, in the sense that Davidson claims, disquotational T-sentences are interpretive translations. In fact, we can infer that they are not just one kind of interpretive translation, but also the most basic kind. For what an interpreter aims to discover, using non-disquotational T-sentences, is of course the meaning of sentences in a foreign language, i.e., what speakers of the language understand in knowing that language. But since speakers of that language do not need to use the interpreter’s language to understand their own, what the interpreter is trying to discover is best represented by a disquotational truth theory. Moreover, the interpreter’s use of a truth theory depends, of course, on his understanding of his own sentences, and hence on a disquotational truth theory for his language.

\(^{145}\) Donald Davidson, “First Person Authority”, in Subjective, Intersubjective, Objective.

\(^{146}\) Donald Davidson, “Pursuit of the Concept of Truth”, in Truth, Language and History, 72.
Perhaps it is worth noting that Davidson has often denied that his account characterizes speakers as actually having or using truth theories:

I do not think I have ever conflated the (empirical) question how we actually go about understanding a speaker with the (philosophical) question what is necessary and sufficient for such understanding… I do not think we normally understand what others say by consciously reflecting on the question of what they mean, by appealing to some theory of interpretation, or by summoning up what we take to be relevant evidence.  

On this account of Davidson’s position, a truth theory can represent or model the linguistic competence of a speaker, but that competence need not itself consist in the ability to use or construct such a theory.

It is hard to square these remarks with the many passages in which Davidson seems to clearly assert that actual understanding involves interpretation, i.e., the use of truth theories. In a paper written after the paper containing the passage cited above, Davidson writes, “understanding, to my mind, is always a matter not only of interpretation but of translation, since we can never assume we mean the same thing by our words that our partners in discussion mean”. If understanding is always a matter of interpretation, and translation, how can it be that truth theories are merely abstract representations of a speaker’s competence? Are we to suppose that Davidson means, instead, that understanding can always be represented as a kind of interpretation or translation? If so, he could have said that. But more to the point, this thesis could not be supported by the claim that “we can never assume we mean the same thing by our words” as do others. This is an empirical claim about our actual epistemic or semantic situation, not a claim about the abstract structure of linguistic understanding. It does not imply anything about how such understanding can be abstractly represented, but it does imply that all actual understanding involves interpretation.

---

Finally, in yet another later paper, and after saying that “almost from the start I held that those who use language do not normally have a [Tarskian truth] theory”, he distinguishes between knowledge of language and merely practical ability as follows:

Knowing a language is, in some respects, like knowing how to ride a bicycle. In both cases, as Rorty points out, we talk of knowing how, and in neither case is it necessary or common to know a theory that explains what we do. But there are also striking differences. There are endless things a speaker or interpreter must know: the truth conditions a hearer will probably take her utterances to have, the truth conditions that most of the sentences she hears will have, relations of entailment, contradiction and evidential support among sentences… Bicycle riding requires no propositional knowledge at all. Davidson is right; this does seem to be a key difference between knowing a language and knowing how to ride a bicycle. In understanding the sentence “Snow is white”, it does seem that I have a kind of propositional knowledge—I know that the sentence means what it does. Or so it seems. And it is that knowledge that we want an account of understanding and meaning to somehow characterize.

Linguistic understanding, in other words, is presumably not simply a matter of behavior or dispositions to behavior, i.e., of some totally non-cognitive skill. But this seems to be at odds with Davidson’s denial that his account of truth-theoretic radical interpretation is purely “philosophical” and has no bearing on the empirical, psychological mechanism of understanding. If knowledge of a language involves “ propositional knowledge”, it would seem that it involves something much like a theory—in some loose sense of the term, perhaps. If the knowledge required is, in particular, knowledge of “the truth conditions a hearer will probably take her utterances to have, the truth conditions that most of the sentences she hears will have, relations of entailment, contradiction”, then knowledge of a Tarskian truth theory, in particular, would seem to be precisely what is required.

It is hard to avoid the conclusion that, despite Davidson’s disclaimers, his position does require that actual understanding is truth theoretic interpretation. But in any case,

---

150 Ibid., 321.
these complications do not affect the main points I have tried to establish in this section. For even if we suppose that linguistic understanding is only represented by an account of (radical) truth theoretic interpretation, it is still true that, for Davidson, this is an exhaustive representation of understanding, and that truth theories—in particular, disquotational truth theories—are supposed to explain meaning and understanding.

VI

Let us now consider how the view of linguistic meaning and understanding under discussion bears on the objection advanced earlier in this chapter to Davidson’s argument from Convention T. The argument, on my reading, was for the conclusion that if two sentences, \( s \) and \( t \), are both in the extensions of a predicate of the form “true in \( L \)”, there is some more general predicate of that form that applies to the extensions of both of the first two. There must then be a truth theory containing the more general predicate, which could be used to translate both \( s \) and \( t \) into some single language. We could say that the theory effectively reveals these sentences to belong to a single language, and this would seem to undermine the idea that \( s \) and \( t \) are associated with divergent schemes.

I objected to this that if the argument is directed at its proper target, it should concern the possibility of understanding or interpretation, and not merely translation, since mere disquotations could be counted as translations by the minimal criteria for translation suggested by Convention T. Because disquotation involves the application of a truth predicate to the quoted sentences by means of a meta-language, there is a sense in which it involves translation.

Disquotational translation, however, does not seem to be sufficient for understanding or interpretation. Davidson’s argument thus rests on a criterion for identity of conceptual scheme so weak that understanding or interpretation \emph{within} a single scheme, so construed, might be just as impossible as understanding across schemes. We should say, instead, that if two languages are associated with the same conceptual scheme, it is possible not merely to translate both into the language of a (single) disquotational truth theory, but also into the language of a (single) truth theory that provides interpretive translations.
In the last section, however, we saw that, for Davidson, the existence of a truth theory is *constitutive* of the meaning of expressions in a language: the nature of meaning, i.e., what it is for language to be meaningful, is exhaustively characterized by the totality of (empirically adequate) truth theories. There is no further dimension of the meaning of expressions in a language over and above the content of a truth theory for that language. Now, as we’ve seen, a truth theory for a language does not have to use sentences other than the disquotations of the object-language sentences to state truth conditions for those sentences. For any language, there will be a set of empirically adequate truth theories that represent the same semantic facts; one of these will be a disquotational truth theory. Therefore, there is nothing that a non-disquotational truth theory says about a language that is not also said by a disquotational truth theory.

Of course, a disquotational truth theory will not be useful to an interpreter whose language is not the object language of the theory. In this sense, the non-disquotational theories have contents that exceed that of the disquotational theory. But this is a fact about the interpreter’s epistemic situation, not about the nature of meaning, i.e., the fact that the language is meaningful. The property of being meaningful is constituted by what is invariant across truth theories for that language, and the interpreter’s language, whatever it may be, is not an invariant feature of all those theories. Moreover, as we’ve seen, disquotational truth theories have semantic and epistemic priority over non-disquotational truth theories, since they reflect what speakers of a language know in understanding their own languages. Other truth theories reflect this knowledge indirectly, by way of the knowledge that a speaker of some other language would have in coming to understand the first.

The key point here is that, on Davidson’s view, the *direction of explanation* is from truth theories, which exhaustively characterize meaning, to understanding, which Davidson identifies with interpretation. For a language to be intelligible or interpretable is, of course, for that language to be meaningful. This is all that is required for the conclusion that the language could, in principle, be understood. (Again, we are ignoring the many easily imaginable ways in which a language could fail to be intelligible or interpretable to some particular kind of speaker, which are not relevant.) But the fact that a language is meaningful is just the fact that there is a Tarskian truth theory applicable to
the sentences of that language. What is understandable is explained by appeal to what is meaningful, and what is meaningful—the nature of linguistic meaning—is explained by appeal to the idea of a truth theory.

There is no requirement that truth theories be non-disquotational; indeed, as we’ve seen, non-disquotational truth theories are needed only because of an interpreter’s ignorance: they reflect the facts of meaning less directly than disquotational truth theories. So the objection raised earlier fails because it presupposes that, once a Tarskian truth predicate has been defined for a language, there is some further dimension of meaning that has not been captured, and which might result in an incommensurability between sentences of the language, or across languages.

The objection could be expressed by asking how we can know that a pair of sentences, \( s \) and \( t \)—which, let us suppose, might be incommensurable—could both be understood, by a single interpreter, merely in virtue of the fact that a disquotational truth theory is applicable to their conjunction. The answer would appear to be that there is no other way in which any sentence is ever understood: the fact that there is a disquotational truth theory for \( s \) and \( t \) is constitutive of the intelligibility of both, and of the intelligibility of their conjunction; since the existence of such a theory is constitutive of the intelligibility of those sentences, it is, of course, sufficient for their intelligibility. On this view, it is not coherent to suppose that, despite the fact that such a theory is possible, no one could understand its object language.

VII

If Davidson’s account of the nature of linguistic meaning and understanding were correct, the objection advanced earlier could be blocked in the way I’ve explained just now. But this account is incoherent. The reason is simple: the very concept of interpretation is not intelligible except if it is taken to depend on the concept of a prior, non-interpretive form of understanding. No one can be intelligibly described as “interpreting” something, in any sense of the term, unless that person is also taken to understand, without interpretation, something else that provides a basis for interpretation.
Corresponding to this distinction between interpretive and non-interpretive understanding, I will argue, there is a distinction between truth in a language or truth relative to a language or system of concepts, on the one hand, and truth \textit{simpliciter}. Nothing can be intelligibly described as true “in” some language unless something else is taken to be simply true, “in” or “relative to” nothing. These considerations are enough to dismantle Davidson’s argument, and to indicate how one might, seemingly, imagine that a sentence or theory is true without committing to any view as to the possibility of translation or interpretation.

Consider a situation in which interpretation, in the normal sense of the term, is clearly required. Suppose that, as in our earlier example, an English interpreter comes to understand a German sentence by means of the T-sentence

\[(T1) \text{“Schnee ist weiss” is true in German if and only if snow is white.}\]

What exactly did he come to know? Not simply that “Schnee ist weiss” is true in German if and only if snow is white. That is what a monolingual German speaker would know in understanding “Schnee ist weiss” (and knowing that he speaks German), whereas the interpreter acquires that knowledge, of course, by means of his own language. He comes to know that

(i) “Schnee ist weiss” is true in German if and only if “Snow is white” is true in English,

but this knowledge serves as an interpretation of “Schnee ist weiss” only because he also knows that

(ii) “Snow is white” is true in English if and only if snow is white,

where sentence (ii) characterizes his understanding of his own language, which he uses to interpret the other. This is obvious, and is, of course, a key feature of Davidson’s account of interpretation. The interpreter, writes Davidson, aims to produce “a characterization of truth-for-the-alien which yields, so far as possible, a mapping of sentences held true (or false) by the alien on to sentences held true (or false) by the linguist”.\textsuperscript{151}

What the interpreter counts as truth-for-the-speaker is defined by what he himself takes to be true, which requires that he consider the truth conditions of his own sentences. The interpreter thus correlates two sentences, one that he already understands and takes to

\textsuperscript{151} “Truth and Meaning”, 27.
be true, given the circumstances, and some other that he does not yet understand, but also
takes to be true. What he ascribes to the alien sentence is thus not merely the property of
truth (in the alien’s language), but more precisely the property of having the \textit{same truth}
\textit{condition} as another certain, which he takes to be true.

So interpretation, in this paradigmatic case, involves taking two sentences to have
the same truth condition. Since the interpreter knows what that condition is, by observing
some seemingly obvious and salient features of the shared environment, he can then
hypothesize that the alien sentence means the same as some sentence that he would be
disposed to assent to, given the circumstances. But now consider the question of how the
interpreter is to understand his own sentence “Snow is white”. We can suppose that his
understanding of the sentence consists in knowledge of its truth condition, as suggested
above—that is, he knows that

\begin{quote}
(T2) “Snow is white” is true in English if and only if snow is white.
\end{quote}

But in this case, it can hardly be that he \textit{interprets} the quoted sentence in the same way as
he interprets “Schnee ist weiss”. That is, his understanding of the sentence “Snow is
white” could not possibly consist in the knowledge that

\begin{enumerate}
\item “Snow is white” is true in English if and only if snow is white
\item (i) “Snow is white” is true in English if and only if snow is white
\item \textit{and also} the knowledge that
\item (ii) “Snow is white” is true in Meta-English if and only if snow is white.
\end{enumerate}

It would be absurd to suppose that one could interpret a sentence that one does not
understand by using that same sentence to state its own truth condition. But there are
other, unattractive consequences of this suggestion, which I take to be decisive reasons to
reject the idea that all understanding is “interpretation” in any univocal sense of the term.

One consequence is a psychologically impossible and, in any case, non-
explanatory infinite regress of interpretation. If the interpreter does not understand,
without interpretation, the sentence “Snow is white”, which he uses to interpret the
sentence “Schnee ist weiss”, he will need to interpret “Snow is white”, i.e., to state its
truth condition by means of another sentence. At each stage of the supposed
interpretation, we can ask how he understands the sentence that states a truth condition.
If the answer is always that he interprets that sentence, he cannot in fact be \textit{interpreting}
anything at all. For how could someone interpret a sentence on the basis of some
(infinite) series of others sentences that he also needs to interpret, and hence does not understand? But, of course, if he understands any of these sentences, and uses that understanding to interpret the others, he could just as well have understood the first, English occurrence of “Snow is white”. The present suggestion is therefore either incoherent or false.

Finally, we should recall that interpretation requires not merely that sentences be assigned truth conditions, but that some of them, at the outset, are taken to be true. Unless the interpreter can discover or assume that some of the speaker’s sentences are true under just those conditions that obtain, in the interpreter’s view, he will never be able to assign truth conditions to any of the speaker’s sentences. What is the property the interpreter ascribes to a speaker’s sentence, on this view, when he takes it to be true? If all understanding is interpretation, he takes the speaker’s sentence to be true in that he takes it to have the same truth condition as a sentence of his own, which he takes to be true. But his understanding of his own sentence consists in his taking it to have the same truth condition of some third (true) sentence, and so on ad infinitum. At no point does he reach a sentence that he simply understands, and takes to say something that describes the world as it is—e.g., that says that snow is white.

What the interpreter ascribes to the speaker’s sentence is thus a purely infra-linguistic relation, the relation of having the same truth condition as some (infinite) number of other sentences. If we wish to maintain that he did in fact attribute truth to the German sentence, we will be forced to deny, incoherently, the truism that truth depends on how the world is. The property that the interpreter ascribes to a speaker’s sentences, on this view, is clearly not truth at all. We could then summarize this last objection by saying that if all understanding is interpretation, there is no interpretation. The assumption that all understanding is interpretation implies that interpreters cannot ascribe truth to sentences, their own or anyone else’s, but at most the property of having the same truth condition as some other sentence. (How anyone could ascribe that property without being able to ascribe truth is, of course, equally hard to imagine.) But interpretation requires the ascription of truth, for only if truth is ascribed can the sameness of truth conditions be discovered; therefore, if all understanding is interpretation, there is no interpretation.
It is not coherent, then, to suppose that all understanding is truth-theoretic interpretation. A speaker’s understanding of his own utterances, at least, cannot be the same kind of understanding that a radical interpreter achieves regarding a foreign language. But why does this matter? After all, we noted earlier that Davidson recognizes a distinction between a speaker’s knowledge of his own meaning and the knowledge he can have of what other people mean by their words. He holds the best and ultimate account a speaker can give of what he means by his own words will consist in a series of disquotational T-sentences. This would seem to be a different way of saying that, as I’ve argued, a speaker cannot interpret his own words in just the same way he interprets other people’s words—for to interpret the sentence in that way would be, presumably, to produce a non-disquotational T-sentence that characterizes the truth condition of the quoted sentence. But it is still true, on this view, that the basic structure of understanding is characterized by a truth theory for a speaker’s language.

Now this may be right, in the following sense. If we doubt that a speaker expresses knowledge, which is normally authoritative, of the meaning of his own utterance of some sentence $s$ by saying something of the form “My utterance of ‘$s$’ is true (in my language) if and only if $s’$, no further interpretation of $s$ will be an improvement. If he is not assumed even to understand the disquoted meta-language sentence $s$, he can’t be assumed to understand any non-disquotational meta-language sentence that might take the place of $s$. For how could he know that the other sentence means the same as the disquoted sentence, if he doesn’t already understand that latter sentence? Thus, we cannot make sense of the idea that a speaker “improves” on disquotation of his own sentences, for the “improvements” could not be evidence of meaning or understanding unless his utterances of disquotational T-sentences are taken to be stronger evidence of that kind.

This may explain why a speaker’s understanding of his own utterances must be characterized, ultimately, by disquotation, but it does not explain why that kind of understanding should be considered as a kind of interpretation—i.e., as being essentially
the same kind of knowledge, in any sense, as the knowledge of a language that a radical interpreter acquires. In what sense might this be true? As we’ve seen, it clearly cannot be the case that a speaker normally interprets his own utterances by comparing them with the environment, on the one hand, and some other sentence or utterance that, given the circumstances, he takes to be true. Again, the only answer would seem to be that, as with the understanding of other people’s utterances, a speaker’s self-understanding can be characterized by a truth theory.

Thus, Davidson’s view would seem to be that what makes an activity interpretive, ultimately, is that it can be characterized in this way, i.e., that it somehow involves a use of the concept of truth. Perhaps we should then say that, while “all understanding of the speech of another involves radical interpretation”¹⁵² one’s understanding of one’s own speech merely involves interpretation in this less radical, more basic sense. What distinguishes the two cases would seem to be just that, in radical interpretation, the interpreter does not understand the language to be interpreted, while in self-interpretation, the speaker (i.e., the “interpreter”) does understand that language. But if he understands the language, he does not need to interpret, and indeed, he could not interpret it. This would just be a way of saying that there is some kind of understanding that is not interpretive. And if that is the kind of understanding that a speaker has of his own language, that understanding is not exhaustively represented by a radical interpreter’s knowledge of the language.

Moreover, if a speaker can understand his own utterances in some way other than by interpretation, it is hard to see why he cannot normally understand other people’s utterances in that way. If I take my utterance of “Snow is white” to mean that snow is white, without having to interpret that utterance, why could I not take your utterance of the very same sentence to mean the same thing without interpreting your utterance? Davidson’s answer to this question typically appeals to the possibility of misunderstanding, as in the passage cited earlier:

¹⁵² “Radical Interpretation”, 125.
Understanding, to my mind, is always a matter not only of interpretation but of translation, since we can never assume we mean the same thing by our words that our partners in discussion mean.\footnote{“Gadamer”, 275.}

The point would seem to be that we can never safely assume that others mean by their words what we would mean by those words; this should be a conclusion based on evidence, to be acquired through translation of the other person’s idiolect. Even if two speakers do in fact use language in much the same way, they can only recognize this on the basis of translation.

This argument seems to be invalid. From the fact that we could be wrong about how others use words it does not follow that we are, of course, i.e., that the natural assumption that other people’s utterances of certain sentences mean just what they would if we were to utter those sentences is actually false. But suppose that the assumption is in fact true, most of the time—e.g., your utterance of “Snow is white” does mean, to you, just what my utterance of that same sentence would mean to me. In that case, it is possible that both of us can operate on the assumption that our idiolects are the same and succeed in understanding one another without interpretation or translation. Indeed, there need not be any such assumption in the literal sense of something believed by either of us, consciously or unconsciously; we might simply understand each other’s words, without given any thought at all to the question of whether we are using words in the same ways.

Interpretation could be required, if one of us begins to use words in ways that the other finds puzzling or unclear, and in that case it would be necessary for the other to consider the differences between our idiolects. But this does not suggest that, when no such difficulties arise, both of us must nevertheless be engaged in the same kind of activity, or that we need to find evidence, by way of translation, to warrant a conclusion about the languages we are using. Once we grant that there must be some kind of non-interpretive understanding, in the case of self-understanding, it is natural to suppose that this same kind of understanding is also possible in other cases. The fact that, in these other cases, interpretation is always a possibility does not imply that it is always going on
even when neither speaker has any reason to treat the other’s utterances any differently from his own.

Given these considerations, the defence of Davidson’s position regarding Convention T suggested earlier fails. The direction of explanation cannot be from truth theories to understanding, by way of interpretation, but must be, instead, from some more basic, non-interpretive form of understanding to interpretation. No one could use a truth theory to interpret another person’s utterances, or to articulate his understanding of his own, unless he already understands, without interpretation, the language of the theory. Since understanding is prior, in the order of explanation, to interpretation or truth theories, the mere existence of a truth theory does not guarantee that the language or languages in its scope are intelligible to anyone: the fact that each of two true sentences $s$ and $t$ are in the scope a single disquotational truth theory does not imply that $s$ and $t$ are jointly intelligible, because it does not imply that the meta-language sentence that states the truth condition of their conjunction is intelligible. In order for that conclusion to follow, we would need the assumption that that sentence can be understood, in some non-interpretive way. But that assumption would beg the question, of course, against the suggestion that $s$ and $t$ may belong to divergent conceptual schemes.

IX

The natural and inevitable distinction between interpretation, which involves coming to understand something on the basis of something that is already understood, and understanding per se, which does not, breaks the conceptual connection that Davidson claims to discover between the notions of truth, interpretation and translation. We can expand on this point, on the basis of Davidson’s contention that to understand a language is to know the truth conditions of sentences. Indeed, the argument I want to make can be based on the even weaker assumption that such knowledge is a necessary condition for understanding a language.

On this assumption, if a speaker understands the sentence “Snow is white”, he must know something that can be represented by a T-sentence. He knows that

(Ts) “Snow is white” is true (in English) if and only if snow is white.
But, as we’ve seen, his understanding of the second occurrence of “Snow is white” cannot depend on or consist in some further knowledge of how to translate or interpret that sentence by means of another—whether the other is a different sentence altogether, or simply a disquotation. He simply knows that “Snow is white” is true if and only if snow is white, which is to say that he knows that the sentence is related in a certain way to a (possible) condition of the extra-linguistic world.

This kind of non-interpretive understanding can be represented by a T-sentence, as in the above example. But a T-sentence can also be used to represent an interpreter’s understanding of a sentence in someone else’s language. What I want to emphasize here is that these are in fact two very different ways of construing T-sentences, or truth theories more generally. On the first kind of construal, a T-sentence is taken to characterize a relation between a sentence and some (possible) condition of the world, and the truth relevant relation so characterized will not be relative to a language—not, at least, in the same sense as it would be if the sentence were under interpretation or translation. Let us call this a representational construal of a T-sentence: the sentence on the right hand side of the biconditional is taken simply to represent the world as being a certain way.

On the other hand, T-sentences can also be construed as translations, and hence as a basis for interpretation. On the translational construal, a T-sentence is taken to indicate a speaker’s knowledge of a relation between sentences, i.e., the relation of having the same truth condition, and not simply between one sentence and the world. Again, a speaker who translates or interprets “Schnee ist weiss” by means of his own sentence “Snow is white” must know, first, that “Snow is white” is true if and only if snow is white, and, second, that “Schnee ist weiss” is true if and only if “Snow is white” is true.

The representational construal is basic, both semantically and epistemologically. For, as in the example just given, a translational construal of a T-sentence rests on an implicit appeal to more information than is actually given by the T-sentence. The sentence “‘Schnee ist weiss’ is true in German if and only if snow is white” does not say anything about the sentence “Snow is white”, but merely about the relation between another sentence and the color of snow. So in order to construe a T-sentence as characterizing the knowledge that a translator or interpreter has, we must presuppose that
he also knows something about the meta-language sentence, which would then have to be expressed by a representational T-sentence. This is what reflects the fact that the interpreter is using one language to understand another. When, instead, we construe a T-sentence representationally, we presuppose nothing about translation or interpretation.

In later writings, at least, Davidson allows that there is a concept of truth prior to the Tarskian concept of truth in a language:

We have to wonder how we know that it is some single concept which Tarski indicates how to define for each of a number of well-behaved languages…Clearly we cannot recognize that such a predicate [i.e., a predicate of the form “true in L”] is a truth predicate unless we already grasp the (undefined) general concept of truth.  

Davidson does not explain how we recognize Tarskian truth predicates to be genuine truth predicates. But clearly this recognition is based the representational aspect of T-sentences, which, unlike the translational aspect, reflects the “the simple concept of being true” (VICS, 193): we recognize sentences of the form “‘s’ is true if and only if s” to be trivially true, as Davidson says, because they express the truism that whether a sentence is true depends on whether the world is as the sentence says. These sentences specify what it is for particular sentences to be true, in a way that does not normally concern their relation to other sentences, but to how things are in the world.

I would suggest that this is what explains why Convention T seems to us to be an acceptable constraint on theories of truth: such theories must at least respect the truism that whether a sentence is true depends on whether the world is as the sentence represents or describes it. Just as obviously, the translational construal or aspect of T-sentences plays no role in explaining the intuitive attractiveness of Convention T. Rather, once we accept this constraint on theories of truth, we can generalize the constraint on the basis of the fact that the same truths can be expressed in different language: a meta-language that translates the object-language, if there is such a meta-language, can also serve to establish the representational relation between sentences and truth conditions.

In light of these considerations, Davidson’s account of the relation between the concept of truth and Convention T seems perverse. According to Davidson, “Convention

---

154 Donald Davidson, “Truth Rehabilitated”, in Rorty and His Critics, 69-70.
T suggests, though it cannot state, an important feature common to all the specialized concepts of truth [i.e., all the instances of truth in L] … by making essential use of the notion of translation” (VICS, 195). He means that *translatability* is the feature common to all the “specialized concepts of truth”, even though, absurdly, this will be in many cases the totally vacuous form of “translation” involved in mere disquotation. At the same time, the *representational* aspect of Convention T, which is all that is actually common to all those concepts, and which is so obviously what accounts for the intuitive correctness of Convention T as a constraint on theories of truth, is entirely ignored.

So, according to Davidson, we are to regard all T-sentences as having a common property that they do not have, except in a highly strained and legalistic sense that is obviously irrelevant to the issue of interpretation or understanding. But we are to overlook the vastly more obvious and important common property: the property of being concepts of truth, concepts that apply to all and only those sentences that describe or represent the world as it is. This puzzling account of the relation between the general concept of truth and the Tarskian truth concepts, which says nothing at all about the dependence of truth on how things are in the world, is supposed to motivated by an intuition concerning “how the concept of truth is used”.

It is also worth noting that the argument from Convention T is invalid. Davidson infers that the concepts of truth and translation are inter-dependent from the premise that Tarskian truth theories make “essential use of the notion of translation”. Instances of the schema “s is true in L if and only if p” are translations, in a certain sense, since the sentence that replaces “p” translates the sentence that replaces “s”. From this, Davidson appears to infer that our understanding of instances of that schema essentially involves the concept of translation. So, the fact that a kind of translation is used in stating truth conditions for a language is supposed to imply that the concept of truth characterized in this way involves the concept of translation. But this doesn’t follow, any more than the fact that we must use words to refer to snow implies that the concept of snow is the concept of a word.

What Davidson would need for his conclusion that the concepts of truth and translation are inter-dependent would be, not merely the premise that T-sentences use translation, in some sense, but the premise that in construing a T-sentence as a
characterization of the truth condition of the quoted sentence, we are construing it as a translation of that sentence. Or, in other words, he would need the premise that in construing a T-sentence representationally, and hence making use of translation, we are also conceiving of the T-sentence as a translation. So, for instance, suppose that in conceiving of the T-sentence

\[(T1) \text{“Snow is white” is true if and only if snow is white}\]
as a specification of what it would be for that sentence to be true, one must also conceive of the second occurrence of “Snow is white” as a translation of the first: one must, that is, be aware of the fact that both occurrences have the same truth condition or meaning. This doesn’t seem to be true. Rather, it seems entirely possible that one might take (T1) simply as a specification of the condition under which the mentioned sentence is (or would be) true, without giving any thought to the sentence used to specify that condition (i.e., the second occurrence of “Snow is white”).

Certainly it is not generally true that every property of an expression figures in the meaning that one takes the expression to have. I can take the words “Donald Davidson” to refer to Donald Davidson without thereby having to conceive of them also as containing occurrences of the letter “d”, or as both having more than three letters. Why then should we not be able to take a sentence like (T1) as meaning simply that sentence “Snow is white” is true if and only if snow is white, while ignoring the fact that the second occurrence of the sentence translates, trivially, the first occurrence? At the very least, some further argument is needed here for the conclusion that translation is not merely used in stating the truth conditions of sentences, but that the concept of translation is or must be used in conceiving of T-sentences as statements of truth conditions.

\[X\]

Let us take stock of the main conclusions reached in this chapter with regard to Davidson’s argument from Convention T. First, we noted some prima facie grounds for doubt concerning the premises of Davidson’s argument. Convention T does not in fact make any reference to English, or to any particular language. So even if Convention T were taken to impose constraints on the concept of truth, and not merely the Tarskian
concept of truth in a language, it would not rule out the possibility of a true theory that
could not be translated into our language. At most, we could conclude that the theory
must be translatable into some language or other, which might also be incommensurable
with ours. So, on this reading, the argument fails.

Second, we considered a more abstract reading of Davidson’s argument, which
rests on the conceptual dependence of predicates of the form “true in L” on meta-
languages that can be used to translate sentences to which those predicates apply. We
saw that this more abstract reading establishes that divergent schemes, on the present
account, would at least be subject to a shared truth predicate, and hence to translation into
a single meta-language. My objection to this latter argument was that the kind of
translation required by Convention T is not sufficient to ensure the possibility of
interpretation or understanding, since merely disquotational translation would be enough.
This objection would seem to be a decisive argument against Davidson’s view that the
constraints imposed by Convention T on truth predicates are sufficient to rule out the
possibility of a difference in conceptual scheme.

I then argued that, on Davidson’s view of linguistic meaning and understanding,
Convention T would in fact seem to be sufficient to rule out that possibility, since it
implies that the direction of explanation must be from disquotational instances of the T-
schema to meaning and understanding. This leaves no room for any intelligible doubt as
to whether a language could be understandable despite being the object of a truth theory.
In response to this possible defense of Davidson’s argument from Convention T, I have
set out what I take to be overwhelming reasons to reject any such view of linguistic
meaning and understanding. Davidson’s position is not merely counter-intuitive, but
involves a reversal of the actual conceptual priorities. Understanding and meaning are
supposed to be dependent on interpretation, and truth is supposed to be dependent on the
translational or interpretive concept of truth in a language. This leads to unacceptable
consequences.

In fact, it is the concepts of interpretation and truth in a language that depend on
the prior concepts of meaning, understanding and truth. While the concepts of truth and
*meaning* are essentially linked, since a sentence cannot be true, of course, unless it means
something, i.e., unless it says that the world is a certain way, a sentence can be
meaningful by being understood, without interpretation. This distinction between meaning and understanding, on the one hand, and interpretation or translation, on the other, severs the connection Davidson purports to find between the concept of truth and the concept of translation, or, more precisely, interpretive translation.

For these reasons, the mere existence of a disquotational truth theory for a language is not sufficient for the possibility of understanding across languages, which is what Davidson’s argument is supposed to establish. Certainly if we suppose that a theory is true, we are committed to the claim that there is a truth theory for that language, since it must at least be possible to use the sentences of the theory to state the truth conditions of those sentences. But this commitment concerns only the representational construal of T-sentences: since the sentences in question are true, they must represent or describe the world as it is. This is implies, of course, that they are meaningful. But they might be meaningful simply by being understood by speakers of the language, and not by being interpreted or interpretable by means of any other language. A disquotational truth theory would characterize that understanding by relating sentences to their truth conditions, not to other sentences in another language.

Since the representational and translational construals of T-sentences are different, and only the former construal is implied by the claim that a sentence or theory is true, no conclusions can be drawn concerning the possibility of interpreting those true sentences by means of our language, or any other. Indeed, it might even be that, as would be required for a difference in conceptual scheme, the understanding possessed by speakers of that language would somehow prevent an understanding of the meanings of expressions in our language, or vice versa. Given only what we’ve seen thus far, it would then seem to be possible to characterize truth, and hence language, in a way that does not presuppose or imply the possibility of interpretation across languages. We need only suppose that both of two groups of sentences represent or describe the world as it is; these may not be sentences that some logically possible thinker could jointly understand or believe. This would then provide a basis, at least, for a theory of conceptual schemes.

However, there are other, more natural and obvious reasons to doubt that this suggestion is intelligible. Suppose that two sentences, s and t, are both true, in the sense that they both represent or describe the world as it is. If these true sentences are
associated with divergent schemes, it must be impossible, in principle, to understand both—impossible to conceive of the world as being the way that $s$ represents or describes it and also to conceive of it as being the way that $t$ represents or describes it. The impossibility here is not psychological, but rather conceptual: it is one’s understanding of $s$ that prevents an understanding of $t$, and vice versa.

This does not seem to be a coherent proposal. If it is impossible in the relevant sense to describe or conceive of the world as being such that both sentences are true, the only explanation of this fact would seem to be that the world simply could not be as both sentences represent or describe it. The impossibility here is supposed to be conceptual, hence akin to the impossibility of conceiving of something as being both round and square, or both red and green. But it is hard to see how this kind of conceptual incompatibility could be explained except, ultimately, by some inconsistency between the contents of sentences, or between the constitutive principles of the concepts in question. Of course, this would be to reject the initial suggestion that both $s$ and $t$ are true. But on the other hand, if there is no inconsistency between the contents of $s$ and $t$, they express merely different and complementary truths, and we cannot imagine why it would be impossible, in the radical sense required for a difference in conceptual scheme, to understand or believe both those truths.

Perhaps it might be held that the impossibility of understanding or believing both truths is just a brute fact, which cannot be accounted for by appeal to some fact about the non-conceptual world. As we saw earlier, Carol Rovane proposes that we construe the “exclusionary relation” that is supposed to hold between divergent schemes in this way; she writes that the relation should be considered a kind of restriction on truth-functional relations, so that truth-functional relations can hold only within schemes but not across them. This would imply that claims belonging to alternative schemes would be neither consistent nor inconsistent.\footnote{Carol Rovane, “On the Very Idea of an Ethical Scheme”, 	extit{Iyyun}, 111-112.} Her suggestion is that we must allow for the sake of argument that certain truths stand in no truth-functional relations to others, and then consider whether, given that assumption, we can make sense of the idea of a difference in conceptual scheme. Otherwise, she
thinks, arguments against the scheme idea will beg the question, by presupposing that all truths are consistent, and hence associated with a single scheme.

The problem with this suggestion is that it begs the question in favor of the thesis that there are multiple schemes. The idea that some truths are neither consistent nor inconsistent with others is no more intelligible than the scheme idea itself; indeed, it seems to me that these are one and the same, although perhaps Rovane’s formulation makes it more explicit what is so puzzling about the idea of a difference in conceptual scheme. Again, it seems fair to ask why there are no truth-functional relations between true sentences, and to ask for some account of this “exclusionary relation” that is consistent with our understanding of truth. This simply does not seem to be the kind of phenomenon that has or needs no explanation.

However, our understanding of truth is surely determined, in part, by a grasp of truth-functional relations. To say that $s$ and $t$ are both true but not consistent (or inconsistent) seems to me no more intelligible than the claim that they are both true but their conjunction is not. So it is hard to imagine how an explanation could be given of this special relation between truths that does not violate our understanding of what it is for sentences to be true. And yet, we are owed some kind of explanation, if Rovane’s suggestion is to amount to something more than the unargued assertion that divergent conceptual schemes are possible. Rovane is right to hold that what is needed for an account of differences in conceptual scheme is a restriction on truth-functional relations, but to suppose that this is intelligible seems to be just to suppose that there could be a difference in conceptual scheme—or, at least, I’ve argued, to suppose that something equally in need of explanation is possible.

The chief conclusion of this chapter is that Davidson’s argument against the proposal that conceptual schemes are true but untranslatable theories is unsound, but that, as with his argument against the proposal that schemes are theories that correspond to the same facts, his conclusion can be defended on less controversial and more natural grounds. This concludes my evaluation of Davidson’s arguments against what he takes to be the two main metaphors intended to convey the idea of a total difference in conceptual scheme, i.e., a total failure of (interpretive) translation across languages. We’ve found that, for various reasons, some of which Davidson ignores or would not
accept, these metaphors cannot be elaborated in a way that provides a criterion of language-hood independent of the possibility of interpretation. If we conceive of any two languages as “organizing” or as “fitting” some common entity, in any of the various ways we have considered, it would seem that those languages are also, in principle, interpretable.
CHAPTER 8
ERROR AND INTERPRETATION

Summarizing his discussion of the idea of total incommensurability, Davidson writes that none of the metaphors he has examined provides “a ground for comparison of conceptual schemes”:

It would be a mistake to look further for such a ground if by that we mean something conceived as common to incommensurable schemes. In abandoning this search, we abandon the attempt to make sense of the metaphor of a single space within which each scheme has a position and provides a point of view. (VICS, 195)

What remains to be considered, he thinks, is a “more modest approach”, which “introduces the possibility of making changes and contrasts in conceptual schemes intelligible by reference to the common part” (VICS, 195). The present chapter concerns Davidson’s discussion of this weaker proposal, that there could be partial or local differences in conceptual scheme, and his view that “massive error about the world is simply unintelligible”. These arguments are related, as we’ll see, in that they both rest on Davidson’s view of the methodology of interpretation.

I

Davidson’s suggestion that differences in conceptual scheme could be “made intelligible by reference to the common part” recalls his earlier claim that “we can be clear about breakdowns in translation when they are local enough, for a background of generally successful translation provides what is needed to make the failures intelligible” (VICS, 192). As we noted, such failures of translation do not reflect a failure of intelligibility or understanding. On the contrary, where we are unable to translate some fragment of a foreign language, it may nevertheless be quite easy to understand what is meant by its constituent expressions.

156 Donald Davidson, “The Method of Truth in Metaphysics”, in Inquiries into Truth and Interpretation, 201.
Local or partial failures of translation are intelligible, according to Davidson, only if there is “an ontology common to the two languages, with concepts that individuate the same objects” (VICS, 192). This seems right, if to “make the failures intelligible” is to identify and describe the differences between the two languages that explain the failure of translation. For this is precisely to interpret, or at least to come to understand, the untranslatable fragments of those languages. I take it that this is what Davidson means in saying that partial failures of translation can be made “intelligible by reference to the common part”: the common part consists of the inter-translatable fragments of the two languages, the “background of generally successful translation” between the two. The inter-translatable fragments provide a common ontology, by means of terms (or concepts) “that individuate the same objects”.

In passing, it is worth noting that for this very reason, pace Davidson, “the metaphor of a single space within which each scheme has a position and provides a point of view” seems entirely appropriate to partial failures of translation—as much as it would to total failures of translation, if there are such failures. Different classifications or descriptions of a single set of objects can be intelligibly compared to “points of view” within a “single space”. In other words, partial failures of translation, on this view, are intelligible not only “by reference to the common [i.e., inter-translatable] part” of the languages in question, but also, consequently, by reference to the common reality to which those inter-translatable fragments apply—as in Davidson’s reference to “the same objects”.

If my language divides the color spectrum in two by the predicates “rerange” and “non-rerange”, and you divide it by the predicates “orange” and “non-orange”, the color spectrum (or, for that matter, the concept of a color spectrum, which we share) could be quite reasonably compared to a “single space”, and our different classifications of that spectrum could be compared to different “points of view” from which that object is perceived.

According to this “more modest approach”, then, a language is supposed to be associated with a conceptual scheme different from our own—partially different, at least—if it is partially untranslatable, although it may be fully interpretable. We can “make intelligible” the failure of translation between some fragment of that language and
some fragment of ours, by interpreting the untranslatable fragment by means of our own language, or perhaps simply by learning that fragment and adding to our language.

Why should we view this kind of difference between languages as a (partial) difference in conceptual scheme of any kind? The idea is that, although the alien language is interpretable by means of our own (and vice versa), interpretation reveals conceptual differences that are profound or extensive enough to warrant talk of a difference in conceptual scheme: a sufficient degree of conceptual diversity, without conceptual inaccessibility, might warrant talk of a difference in scheme. For reasons already considered, I think this is confused. However, Davidson appears to think it worth arguing against, since his conclusion is that “given the underlying methodology of interpretation, we could not be in a position to judge that others had concepts or beliefs radically different from our own” (VICS, 195). In the next section, I’ll examine this argument and explain why I think it does not establish anything of interest as concerns the idea of a partial difference in conceptual scheme.

II

The features of the methodology of interpretation that underwrite Davidson’s conclusion can be understood by reviewing Davidson’s theory of radical interpretation, which we discussed in the last chapter. Again, this is a situation in which an interpreter encounters people whose language and beliefs are totally unknown, and must somehow devise “a theory that simultaneously accounts for [propositional] attitudes and interprets speech, and which assumes neither” (VICS, 195). Such a theory is possible if the interpreter can discover at least some of the speaker’s attitudes of assent and dissent, directed to sentences.

Davidson does not explain how these attitudes are to be discerned. Presumably they are recognized on the basis of such evidence as facial expressions, gestures, inflection or tone of voice, together with very general assumptions about the speaker’s interests, emotions, and his cognitive and perceptual abilities. If the speakers we are trying to interpret are human beings, this seems reasonable enough: there are presumably some behaviors, dispositions, emotions and interests common to all human beings. At the
very least, we should expect that the biological similarities between human beings will provide some basis for mutual interpretation.

Davidson allows that “attitudes are indeed involved here”, since the interpreter must take speakers to have beliefs, i.e., to hold true certain sentences, but the fact that the main issue is not begged can be seen from this: if we merely know that someone holds a certain sentence to be true, we know neither what he means by the sentence nor what belief his holding it true represents. His holding the sentence true is thus the vector of two forces: the problem of interpretation is to abstract from the evidence a workable theory of meaning and an acceptable theory of belief. (VICS, 196)

On the basis of behavioral evidence, the interpreter can establish at least that a speaker holds true (or false) certain sentences, even though he doesn’t know what those sentences mean or which beliefs they express. He can then begin to piece together a theory of meaning and belief for an alien language by finding “a systematic correlation between sentences held true [in the interpreter’s language] with sentences held true [in the speaker’s language]” (VICS, 197).

The interpreter establishes this correlation by reference to the objective circumstances in which the speakers of the alien language assent to sentences. If, for instance, he finds that the aliens assent to a certain sentence $s$ whenever he himself would assent to the sentence “It’s raining”, and not otherwise, he will hypothesize that $s$ means that it is raining. So, if we are to interpret a totally alien language, we must assign “to sentences of a speaker conditions of truth that actually obtain (in our own opinion) just when the speaker holds those sentences true” (VICS, 196).

However, while the evidence on which he bases his theory “simultaneously accounts for attitudes and interprets speech” and “assumes neither”, the step from hold true attitudes to truth conditions requires the interpreter to assume that the speaker’s beliefs are to some significant extent the same as his own. Sentences held true are evidence of beliefs; if we take the aliens to hold true various sentences that we believe to be true under the circumstances, this is just to assume or discover that they share many of our beliefs. This is what Davidson somewhat misleadingly calls “charity”. The interpreter reasons that a sentence $s$ should be translated by “It’s raining” because, in his
view, it *is* raining whenever the speaker holds *s* true. Obviously this inference is plausible only if the speaker can be assumed to notice and believe that it’s raining.

More generally, Davidson holds that if the interpreter is to assign truth conditions to the speaker’s utterances, he must assume that the speaker has largely true beliefs about what seem to the interpreter to be the most obvious, salient features of his immediate environment. In this way, the interpretation of speech depends essentially on “a foundation—*some* foundation—in agreement” (VICS, 197). Unless the interpreter assumes that the speaker is, at least in this minimal sense, largely right about the objective world, he will not be able to establish any correlations between the speaker’s utterances and any potential set of truth conditions. Therefore, we could not find that others had “radically different” concepts or beliefs. Of course, we can still find that their concepts or beliefs are *different* from our own, but “we improve the clarity and bite of declarations of difference, whether of scheme or opinion, by enlarging the basis of shared (translatable) language or of shared opinion” (VICS, 197).

So, to summarize this reasoning: Any interpretation of a speaker’s language requires that we assume or discover that his propositional (hold true) attitudes correspond to some significant extent to what we take to be the objective situation, and hence to our own. In a great many cases, we must find that the sentences he holds true are *are* true, by our own standards, at least. This is because, in radical interpretation, there is nothing *other* than the objective situation (or, again, what we take the objective situation to be) that could provide any content to his utterances. If we assume, instead, that the sentences the speaker holds true are typically or always false at the time and place of his assent to the sentence, we have no way of determining what, if anything, he means or believes. But in correlating someone’s utterances with what we take to be the objective situation, we are, of course, also correlating his utterances with our own beliefs and concepts.

It is unclear how these considerations bear on the thesis that some language or system of beliefs might be “radically different from our own”. As we noted at the outset,

---

157 This is not entirely right. For we might, of course, choose instead to interpret the speaker as typically or always denying precisely the proposition that we ourselves would, at that time and place, affirm—that he means to say that it isn’t raining whenever it is raining, that various red things are not red, etc. But this would be an unjustifiable interpretation; for we could offer no explanation of why the speaker believes what he supposedly does. It would be more reasonable to doubt the speaker’s sincerity, or, better, our translation of whatever expression in his language supposed to indicate negation.
the position that Davidson means to attack in these passages is not that there could be
concepts or beliefs so unlike our own that we could not understand the language
associated with those concepts or beliefs. Rather, “changes and contrasts in conceptual
schemes” are supposed to be “intelligible by reference to the common part”. A “radically
different” system of concepts or beliefs, on this view, is intelligible, and on the basis,
moreover, of some shared set of beliefs and concepts. It is just said to be very different
from our own.

The claim that there could or could not be “radical” differences in this sense is
simply too vague to evaluate, it seems to me. Let us suppose that there is some limit
beyond which the basis of agreement needed for interpretation would be so slight that we
would not be able to make much sense of some alien language. Just how different could
a speaker’s concepts and beliefs be while still being intelligible to us? Davidson’s
argument provides no answer to that question. Moreover, the position he has described
does not itself involve any claim that some particular degree of intelligible difference is
possible, however that might be measured. So there does not seem be any substantive
position here to discuss, whether Davidson’s own or that of the dialectical opponent with
whom he takes himself to be disagreeing.

Of course, if by “radical” differences, Davidson means differences so great as to
preclude interpretation, it is tautologically true that speaker and interpreter could not
differ radically: if we are able to interpreter a speaker, his concepts are not so different
from ours as to prevent interpretation. But surely Davidson doesn’t mean to argue for a
claim that would be true regardless of what the methodology of interpretation might be,
and which would, in any case, be granted by proponents of partial incommensurability.

Ultimately, however, it seems to me that the question of precisely how much in
the way of shared concepts or beliefs is required for interpretation is irrelevant to the
question of whether there could be conceptual schemes partially different from ours. The
very idea of incommensurability—whatever, if anything, it may amount to—implies a
failure of interpretation across “incommensurable” languages, and we have already seen
why there is little philosophical interest in other kinds of differences. Earlier in “On the
Very Idea of a Conceptual Scheme”, Davidson rightly dismisses the idea that cases where
“what comes easily in one language may come hard in another” can be taken as evidence of a difference in scheme:

Examples like these, impressive as they occasionally are, are not so extreme but that the changes and the contrasts can be explained and described using the equipment of a single language. Whorf, wanting to demonstrate that Hopi incorporates a metaphysics so alien to ours that Hopi and English cannot, as he puts it, ‘be calibrated’, uses English to convey the contents of sample Hopi sentences. Kuhn is brilliant at saying what things were like before the revolution using—what else?—our post-revolutionary idiom. Quines gives us a feel for the ‘pre-individuative phase in the evolution of our conceptual scheme’, while Bergson tells us where we can go to get a view of a mountain undistorted by one or another provincial perspective. (VICS, 184)

These authors appear to hold that a language that we can interpret by means of our own is nevertheless associated with a different, i.e., incommensurable and hence uninterpretable, conceptual scheme. Since the very idea of an alternative conceptual scheme is the idea of a system of beliefs or concepts so unlike our own that we cannot understand it, this position is incoherent.

If Whorf, for example, has actually succeeded in using English to explain the meaning of various Hopi sentences, those sentences cannot mean something that is not expressible in English. On the other hand, if Whorf doubts that he has ever fully or accurately captured the meaning of Hopi expressions by means of English, he should also doubt whether he has discovered that Hopi metaphysics is radically unlike ours, the only evidence for this hypothesis being his interpretation of Hopi by means of English.

How then does the present account of conceptual schemes, taken to be “radically” different but intelligible views of the world, differ from the self-defeating kind of account Davidson refutes in the passage cited just now? Again, the idea is supposed to be that we might make “changes and contrasts in conceptual schemes intelligible by reference to the common part”. Davidson does stipulate that the parts of some alien language associated with a different scheme are untranslatable (VICS, 185), and perhaps—depending on how exactly translatability is understood—this distinguishes the proposal under discussion from those Davidson has earlier dismissed. But the basic position is the same: although
we can understand the concepts and beliefs in question, they are, supposedly, associated with an incommensurable and (partially) unintelligible system of concepts or beliefs.

This position is either incoherent or philosophically uninteresting. (Of course, it may well be anthropologically or historically or psychologically interesting.) It has already been discredited merely by his sensible observation that any “changes and contrasts” that can be “explained and described using the equipment of a single language” are, by definition, changes and contrasts within a single conceptual scheme. No additional argument from “the underlying methodology of interpretation” is needed.

Of course, it might be that the Hopi Indians have a “conceptual scheme” unlike our own if that means merely that some of their beliefs and concepts are very different from ours. But, again, this is a trivial truth (if it is a truth), and philosophically insignificant. What we should consider, instead, is the possibility that certain parts of the language or system of belief and concepts of some other group of speakers might be not just very different from ours, but partially uninterpretable or unintelligible. That, presumably, is what it would mean to say that their scheme is partially or locally different from ours. In the next section, we’ll consider how Davidson’s discussion might be taken to bear on that issue.

III

I’ve stressed that, given the course of Davidson’s discussion, a partial or local difference in conceptual scheme should be evident, not where a fully interpretable language contains an untranslatable fragment, but where an otherwise interpretable language contains an uninterpretable fragment. And yet Davidson seems to restrict his attention to the suggestion that some fully interpretable language might be associated with a system of concepts or beliefs “radically different from our own”:

If we choose to translate some alien sentence rejected by its speakers by a sentence to which we are strongly attached on a community basis, we may be tempted to call this a difference in schemes; if we decide to accommodate the evidence in other ways, it may be more natural to speak of a difference of opinion. (VICS, 197)
Davidson’s point here seems to be that there is no way to distinguish between purportedly conceptual and purportedly doxastic differences. Where we might wish to speak of a conceptual difference, we can always “accommodate the evidence in other ways” and conclude that the difference is merely a difference of opinion.

If this were true, it would be devastating for proponents of the scheme idea. But this argument depends essentially on the assumption that the alien language in question is interpretable, and shows merely that interpretable difference between languages or speakers can, in principle, be understood as difference in concepts or, instead, a difference in opinion. The idea of difference in conceptual scheme, however, just is the idea of a language or system of concepts that differs from our own conceptually but not doxastically. Doxastic differences logically require concepts shared by those who disagree: you can’t agree or disagree with my claim that roses are red if you lack the concept of a rose, or the concept of redness.

As was suggested in the last section, what we should consider instead is whether it is intelligible to suppose that a partially interpretable language might contain a fragment that is not interpretable. We can focus this question by considering how this scenario differs from that of total incommensurability. Suppose that some alien language or system of concepts consists of two fragments, one of which is fully interpretable by means of ours, while the other is totally uninterpretable. This would then be an instance, within a single language, of the same kind of incommensurability that would hold between languages associated with divergent schemes. The interpretable fragment is associated with our scheme, while the uninterpretable fragment is associated with some other. But, of course, the mere fact that speakers of the alien language understand both fragments implies that the two are not genuinely incommensurable: it is not impossible, in principle, to understand both. So the idea of a partial difference in scheme cannot be taken to involve the idea that it is impossible, in principle, to understand both fragments of a language that we can partially interpret.

Rather, the suggestion here must be that we could not understand both fragments, despite understanding one of them. It is this, presumably, that Davidson means to deny. This is a key difference, then, between my reading of Davidson’s rejection of the idea of a total difference in conceptual scheme and his rejection of the idea of a partial difference.
in scheme. As concerns the former, I’ve taken his position to be that there could not be two languages or systems of concepts such that merely to understand or make use of one would prevent an understanding of the other—and not that we, with our actual capacities, would be able to understand any logically possible language. In rejecting the idea of a partial difference in scheme, however, Davidson must be taken to endorse the stronger claim that, if we can interpret any part of a language, we can interpret the whole language. Since it is built into this scenario that someone understands both fragments, the controversial claim can then only be that we could not understand both.

Now it might be that Davidson takes himself to be making essentially the same claim as concerns both the idea of a total difference in scheme and the idea of a partial difference: the claim that we could understand any language purportedly associated with an alternative scheme, whether the difference between ours and the other is total or partial. In that case, I disagree with Davidson; I believe that in earlier chapters I’ve given some powerful reasons to reject this claim as concerns the idea of a total difference in scheme. There might well be languages that we could not understand at all, even in principle.

However, as we’ll see, there are special considerations applicable to purported cases of a partial difference in scheme that would warrant the stronger thesis that we must be able to interpret any such language, and not merely that interpretation is logically possible. In any case, it would certainly be consistent to hold this stronger view as concerns the idea of a partial difference in scheme, while holding only the weaker view as concerns the idea of a total difference. What are these special considerations? Although Davidson does not explicitly say so, the feature of “the underlying methodology of interpretation” that is supposed to ensure that we can understand any whole language if we can understand a part of the language is the necessity of a holistic approach in constructing theories of meaning and belief:

If sentences depend for their meaning on their structure, and we understand the meaning of each item in the structure only as an abstraction from the totality of sentences in which it features, then we can give the meaning of any sentence (or word) only by giving the meaning of every sentence (and word) in the language.158

---

158 Donald Davidson, “Truth and Meaning”, in Inquiries into Truth and Interpretation, 21.
If this is right, then no language could be only partially interpretable. For this would be to give the meaning of certain sentences in the language without giving the meaning of every other sentence in the language. To put the point in mental terms, if we can discern any of a speaker’s beliefs or concepts, we can discern all of them.

This reading of Davidson’s position is not clearly supported by his discussion in “On the Very Idea of a Conceptual Scheme”, although it is a central theme in all of his writings on meaning and interpretation. This reading explains Davidson’s view that an argument from the methodology of interpretation bears specifically on the notion of a partially uninterpretable language: by hypothesis, we can interpret some of the alien language, and if holism is true, this implies that we can, at least in principle, interpret all of it. We are to extrapolate from the features of interpretation that are discernable as concerns the interpretable fragment of a language to the conclusion that any other, purportedly uninterpretable fragment of the same language must also be interpretable. This is guaranteed by the assumption, which is built into the present account of a (partial) difference in conceptual scheme, that these are in fact fragments of a single language, and hence holistically inter-connected: the expressions in one fragment are given meaning by their relations to expressions in all the other fragments of the language.

However, the radical thesis of the total holism of meaning and belief is extremely implausible. Certainly it seems that in understanding a speaker’s utterances, one must view him as having a system of beliefs sufficient to support and give content to any particular belief ascribed to him, and this implies a kind of holism:

Someone who can interpret an utterance of the English sentence “The gun is loaded” must have many beliefs, and these beliefs must be much like the beliefs someone must have if he entertains the thought that the gun is loaded. The interpreter must, we may suppose believe that a gun is a weapon, and that it is a more or less enduring physical object. There is probably no definite list of things that must be believed by someone who understands the sentence “The gun is loaded”, but it is necessary that there be endless interlocked beliefs.159

In other words, there is no such thing as a discrete belief, the content or identity of which does not depend on its relations to many other beliefs (even, perhaps, to “endless

159 Donald Davidson, “Thought and Talk”, in Inquiries into Truth and Interpretation, 158.
interlocked beliefs”). So, to ascribe any one belief to someone is to ascribe a system of beliefs. Although, as Davidson notes, there may be many systems that would support any particular ascription. Even if the system of beliefs necessary for the content and identity of any one belief were literally “endless”, i.e., infinite in number, this would still be a more moderate form of holism than the kind expressed in the first passage.

The moderate thesis says that, for any belief, there is some system of other beliefs without which that belief could not what it is. Or, to put the point in linguistic terms, the moderate thesis says that the meaning of any one sentence is determined by its relations to countless others. Let us suppose, for the sake of illustration, that these others are in fact “endless”. The radical thesis, by contrast, says that, for any belief or sentence, the system of other beliefs or sentences that determines the content and identity of that first belief or sentence comprises every belief that a speaker has, or every sentence in his language. These claims are obviously not equivalent, and the first does not imply the second. My belief that the gun is loaded might depend holistically on countless other beliefs that I have—beliefs about weapons, about physical objects, life and death, etc. But even if that is so, I may still have countless other beliefs that play no role in determining the content and identity of that belief. (Moreover, it seems to me that this is just what we would expect, intuitively—that, e.g., my belief that Quine’s philosophy is complex plays no role in determining the content of my belief that a gun is loaded.)

Logically speaking, one could have an infinite number of “endless, interlocked beliefs” about totally disparate subjects that are not holistically connected with each other in the way that total holism would imply. In that case, it would be true that every particular belief is “interlocked” with an infinite number of others, but false that every belief is “interlocked” with every other belief. For example, suppose (absurdly) that some thinker actually has an infinite number of beliefs about arithmetic: for any number \( n \), he believes of \( n \) that \( n + 1 \) is greater than \( n \). There are countless other beliefs, though, that are not among these—beliefs about geography, theology, etc. So it would seem to be possible for me to interpret someone else as having a certain belief, even if this requires ascribing to him countless other beliefs, without ascribing to him all the beliefs that I have or all the beliefs that he has. Moderate holism thus leaves open the possibility that a speaker’s language or system of beliefs might be only partially interpretable to us.
IV

The argument for the thesis of moderate holism seems to stem from Davidson’s view of interpretation, as in the second passage cited above. By contrast, the argument for the more radical thesis, which Davidson does not clearly distinguish from the moderate thesis, stems from the formal requirement that theories of meaning must be compositional. If a truth theory for a language is adequate, it will fix the meaning of an expression only by its global contribution to the truth conditions “the totality of sentences in which it features”. But then, if the meaning of any one sentence depends on the meanings of its constituent expressions, the meaning of any sentence depends on its relations to all the others in the language: “we can give the meaning of any sentence (or word) only by giving the meaning of every sentence (and word) in the language”.

Unlike the more moderate thesis that emerges from reflection on interpretation, this is a requirement of total holism but does it lead to Davidson’s conclusion that we can understand part of a speaker’s language only if we can, in principle, understand the whole? No, because this completely abstract claim about the compositional nature of meaning does not give us any indication of the scope of a workable truth theory for a natural language (assuming such a thing is possible). Perhaps the meaning of every expression is an abstraction from “the totality of sentences in which it features”, but that totality need not be the totality of all the sentences a speaker understands. This suggests that there is a distinction between the holism of beliefs and concepts, on the hand, and the holism of truth theoretical meaning, on the other.

Let us suppose that the extent of holistic dependence between sentences is taken to be the criterion for distinguishing between “languages”, i.e., that a language is just any range of sentences for which a truth theory can be given. If so, it might well be that any speakers will have many different “languages” at his disposal, with the holistic interdependence of sentences being highly localized. But then, on this suggestion, the speaker’s total system of beliefs and concepts can’t be identified with any of the lesser systems of belief associated with any one of those many languages. Rather, a speaker’s

---

total system of beliefs and concepts should be associated with his linguistic *repertoire*, which may comprise any number of disparate languages.

So Davidson’s contention that the meaning of any one term or sentence depends on its relations to all the others in the language, even if it is true, does not establish that it must depend on its relations to all other terms or sentences in a single speaker’s linguistic repertoire, since it does not establish that that repertoire must be construed as a single language in the present sense. Hence the ability to interpret a speaker’s utterances, if these are taken to belong to a single *language*, in this sense, does not imply the ability to understand all of his beliefs and concepts. For that conclusion, we would need the further assumption of total holism concerning beliefs and concepts, which would beg the question.

To suppose that a single speaker has disparate systems of belief, and so, in that sense, speaks many languages, is not to suppose that his beliefs are somehow incommensurable. That would be incoherent. Rather, they must be seen as simply different and complementary. For instance, it might be that, if we had a correct truth theory for English, we would discover that the expression “good” does not make a separable and invariant contribution to all of the complex expressions in which it occurs (“good stew”, “good person”, “good with animals”, etc.). Perhaps the meanings of these expressions have no more in common than the meanings of “Cape of Good Hope” and “Superman’s cape”. In that case, these different semantic roles of “good” would be captured by different truth predicates. Certain (true) sentences containing “good stew” would be in the extension of some restricted predicate—“true in culinary English”, say—while other (true) sentences, containing “good person” would be in the extension of another predicate, “true in moral English”. The mere fact that the word “good” is common to all these expressions does not imply that a single truth theory is appropriate for all of them, and the truth theoretical requirement of total holism does not imply this either, since it does not say that a speaker can have only one language, in the special sense characterized by that constraint.

Of course, it certainly seems that there are *some* expressions common to every “language” in our linguistic repertoire, which would then imply that our repertoire, at least, contains just one language. The sentences “Here are five apples” and “There are
five reasons to believe Davidson’s thesis of meaning holism” appear to contain terms that express the same concepts, not logically simple and unrelated concepts, e.g., the concept five dollars and the concept five reasons. But this appeal to intuition cuts both ways. It does seem that I mean the very same thing by “five” in using either of these sentences. But surely it also seems that someone who has no idea what is meant by the phrase “thesis of meaning holism” can nevertheless understand perfectly what is meant by “Here are five apples”. Both of these intuitions are fairly powerful, so there is no particular reason, at least, to decide on intuitive grounds that total holism is true.

In conclusion, neither the thesis of moderate holism drawn from Davidson’s theory of interpretation nor the holistic constraint on compositional theories of meaning rules out the possibility of a conceptual barrier between certain beliefs and others, all of which are nevertheless beliefs and concepts belonging to a single speaker and expressible by means of a single speaker’s conceptual repertoire. If there were such barriers, certain interpreters might then be unable to understand some of a speaker’s beliefs and concepts but not others. Indeed, it would seem that this is a familiar phenomenon. A ten year old will be able to understand many beliefs and concepts of, say, the world’s smartest cosmologist, but not all of them. The latter’s beliefs about birthdays or elementary arithmetic may be intelligible to such an interpreter, while his beliefs about the origins of the universe may not be.

Of course, it would always be possible to maintain a more radical form of holism by suggesting that even the first set of beliefs and concepts are not really or fully intelligible to this kind of interpreter. We might say that the interpreter understands some of the content of the cosmologist’s belief that today is Suzie’s birthday, or that there are five apples on the table, but not all of its content; or perhaps that he understands all of it in a vague or somehow indeterminate way.

One problem with this suggestion is that it seems to require a dimension of understanding, and hence meaning, that could not play any role in the public use of language. For there could be no evidence, not even in principle, that the interpreter somehow fails to fully grasp what the speaker believes in such cases. (Of course, the fact that he cannot intelligently discuss string theory, say, will not count as evidence unless we presuppose total holism.) This would then be a form of meaning that plays no role in
characterizing the shared, public meaning of sentences, and hence would undermine the Quinean publicity constraint on meaning and understanding: on this view, two speakers could have the very same patterns of assent and dissent with respect to some group of sentences in response to a shared environment, and yet, somehow, what they mean and understand by those sentences could be quite different.

Another problem with this rejoinder is that it seems plainly *ad hoc*. Some degree of holism seems warranted precisely because, when we reflect on any particular belief, as in Davidson’s example of the belief that the gun is loaded, we can see how it is related, in ways that seem to fix its content, to many other beliefs. But we certainly do not come to see, in this way, that that belief is related to every other belief; moreover, it seems *impossible* to imagine this—impossible to imagine even in particular cases how one belief is holistically related to any other arbitrary belief.

This form of meaning and understanding, then, would be indiscernible not only to a radical interpreter, but, seemingly, to speakers introspecting their *own* beliefs and concepts. At least, it does not seem to be something that we can consciously recognize; we would then be forced to the conclusion that these holistic relations are unconsciously known, if they are known at all. Or, if this seems unlikely, would we perhaps have to say that the contents of a speaker’s beliefs and utterances are not known in *any* sense, either to the speaker or to someone who interprets his utterances?

Here we are obviously very far removed from any ordinary conception of what it is to mean or understand something by one’s own words. It seems doubtful that it is really the concept of meaning or understanding that is under discussion. What a speaker means or understands by his utterances should not radically transcend his conscious knowledge. (Of course, what he *implies* by his utterances typically does, and in that sense, one might say that his meaning does as well. But I am assuming here that there is a basic and indispensable notion of meaning that is tied to *understanding*, and which can be divorced from that of implication.) And since, as we’ve seen, the more radical holistic thesis does not follow from the more plausible, moderate thesis or from Davidson’s requirement that theories of meaning be compositional, there is no particular reason to make such bizarre *ad hoc* revisions to our ordinary conception of meaning and belief.
My main conclusions thus far have been two. First, the thesis of moderate holism does not lead to Davidson’s conclusion that all the beliefs and concepts of a speaker must be intelligible to us if any of them are. Second, there does not seem to be any particularly compelling argument for the thesis of total holism, which would ensure that conclusion. In this section, I want to make an additional argument against the argument from the thesis of total holism, if that is indeed the argument that Davidson takes to rule out the possibility of a partial or local difference in conceptual scheme, i.e., the idea of a speaker whose utterances would be only partially interpretable to us.

As is perhaps obvious by now, the thesis of total holism would prove far too much, since it would make substantive disagreement of any kind a conceptual impossibility. Worse, the thesis would seem to be not merely consistent with the possibility of a difference in conceptual scheme, contrary to Davidson’s aims, but in fact to imply that such differences are possible, and indeed that they are ubiquitous. Let us suppose that, as the thesis says, the identity and content of any one belief depends on its relations to every other belief in the system to which it belongs, where the system is taken to comprise all the beliefs of some thinker. What are we then to make of the seemingly obvious fact that any two people will have at least some differences of opinion on some topics? If this is granted, it follows that no interpreter can ever ascribe to a speaker precisely the belief that the speaker actually has, and no two thinkers can ever actually disagree about anything.

For instance, an interpreter may believe that there should be very strict gun control laws, while the speaker believes there should be none. In that case, when the interpreter ascribes to the speaker the belief that the gun is loaded—or rather, a belief that he would call “the belief that the gun is loaded”—he cannot be ascribing to the speaker any belief that the speaker actually has. For as a result of their disagreement about gun control laws, they cannot have beliefs about guns, or anything else, for that matter, with identical contents. If their beliefs and concepts differ in more extensive ways, the situation will presumably be much worse—if the interpreter is, say, Genghis Khan and the speaker is Donald Davidson. So it would seem that each of us could maintain
Davidson’s thesis that no system of beliefs could be less than wholly intelligible to us (or rather, something somewhat like that thesis, since none of us has exactly Davidson’s beliefs) only at the cost of conceding that we have never in fact succeeded in making anyone’s beliefs intelligible to ourselves.

On closer examination, however, this description of the situation is itself incoherent on the assumption of total holism. For we cannot coherently say, of course, that the interpreter believes *that there should be gun control laws* while the speaker rejects that same belief, since both could entertain the proposition that there should be gun control laws only if both had exactly the same beliefs. At most, it is possible for someone to coherently describe the situation by saying that one of these people, either the interpreter or the speaker, has a belief about gun control laws, while the other has some other beliefs that cannot be accurately characterized by the person who is describing the situation. If one were to ascribe any particular beliefs to either of these people, one would have to ascribe to that person all and only one’s own beliefs; but then the difference between that person and the other could not be characterized in any detail, because the other’s beliefs would be indescribable. The thesis of total holism thus leads to seemingly absurd consequences: correct interpretation requires total agreement, and substantive disagreement is not only non-existent, but is non-existent because it is a conceptual impossibility.

What is more puzzling is that, as the preceding discussion suggests, this thesis would also seem to allow for a kind of conceptual difference that has precisely the character of a partial difference in conceptual scheme, or perhaps even a total difference. For if two speakers have different beliefs, and hence concepts, they will differ conceptually in such a way that it is impossible for either to fully understand what the other says, merely because of his own concepts. And if we can allow that far greater differences are possible, as they certainly would appear to be, it would then seem intelligible to suppose that some speakers would be *totally* incapable of understanding what others say. Moreover, for the reasons just given, there could not be a third party who would be able to *describe*, by appeal to some further, more comprehensive system of beliefs, the beliefs and concepts of the first two speakers.
The thesis of total holism would thus seem to provide an account of meaning and belief that is both necessary and sufficient for a theory of differences in conceptual scheme, and indeed a theory of total difference. Interestingly, Davidson has at times argued in this way, although he stops short of this radically relativistic conclusion:

How clear are we that the ancients—some ancients—believed that the earth was flat? *This* earth? Well, this earth of ours is part of the solar system, a system partly identified by the fact that it is a gaggle of large, cool, solid bodies circling around a very large, hot star. If someone believes *none* of this about the earth is it certain it is the earth he is thinking about? Davidson does not answer his rhetorical question, but one wonders what he could possibly say about this that would be both credible and consistent with total holism. Would he say that these ancients actually had no beliefs at all? Or that they believe that something *other* than the earth was flat? If so, what could that have been?

Presumably we cannot say that these people believed nothing; they were not mindless, simply in virtue of not believing what we do about the earth, or, as the case may be, not believing anything at all about that thing. So it would seem that, although we can know that these people had beliefs and concepts, we could not say *which* beliefs and concepts those were, simply because our beliefs and concepts are so different from theirs. Perhaps they had beliefs *like* our beliefs about the earth, in some sense, but not beliefs that we can describe in any more precise way; the contents of their thoughts and utterances are not accessible to us. And no one else could characterize this situation more precisely either, because any characterization would succeed in specifying at most one of the two systems of belief.

Again, this would seem to be precisely what is needed for a difference in conceptual scheme. It is thus quite surprising that Davidson’s holism is commonly taken to be an argument against the possibility of alternative schemes. On the one hand, the moderate form of holism, according to which an ascription of beliefs to a speaker requires that the interpreter be able to ascribe a great many other beliefs, does not ensure that all of the speaker’s beliefs are conceptually accessible to us, or to any particular interpreter. But this argument is often used. Carol Rovane writes that “the pressure of

---

161 “*Thought and Talk*”, 168.
holistic relations”, by which she means moderately holistic relations, “ensure that no belief or concept can stand on its own, without any normative relation at all to other beliefs and concepts”, and calls this “an argument against conceptual schemes”,\(^\text{162}\) i.e., against the idea that some systems of belief could be partially or wholly inaccessible to some interpreters.

If she is inferring from moderate to total holism, this is a logical mistake. As we noted earlier, the fact that no belief can “stand on its own”, but must rather be ascribed as part of a larger system, does not imply that all beliefs of any speaker must be related somehow to all other beliefs of that speaker, or to all beliefs and concepts generally. It is consistent with moderate holism, then, to suppose that some or all of a speaker’s beliefs and concepts might be inaccessible to some particular kind of interpreter. In particular, it might be that some or all of the beliefs and concepts of certain logically possible speakers are inaccessible to us.

On the other hand, the thesis of total holism would secure Davidson’s conclusion that a speaker’s beliefs and concepts are partially intelligible to us only if they are wholly intelligible to us, but at the cost of some very unwelcome and ultimately self-defeating consequences. It would follow, on this view, that there could not be a partially uninterpretable language or system of belief, but also that we could not discover any differences at all between our own beliefs and concepts and those of others. This simply does not seem to be true. In any case, the thesis of total holism does say that no belief can “stand on its own”, but it does not say that all beliefs stand together, i.e., that any system of beliefs is holistically related to all others. On the contrary, total holism implies that any difference at all in belief results in conceptual inaccessibility, as in the example of the ancients’ beliefs about the earth.

So, if total holism is true, then either (i) we are able to interpret others, but only because they (necessarily) have all and only the beliefs that we have, or (ii) there are sometimes differences in belief, and in any such cases neither thinker is able to understand any of the other’s beliefs. These are conclusions are unattractive, to put it mildly. It would surely be preferable to reject the thesis of total holism, which has little intuitive force to begin with. But then we have not found any reason to accept

Davidson’s claim that a speaker’s concepts and beliefs are partially intelligible to us only if they are wholly so.

VI

It is interesting to note that the kind of argument from holism to relativism suggested above appears to be much like an argument that Davidson briefly discusses in “On the Very Idea of a Conceptual Scheme”. Recall that, according to Davidson, one might argue for conceptual relativism on the basis of arguments against the analytic/synthetic distinction. If “meaning, as we might say, is contaminated by theory” (VICS, 187), there is no way to preserve a fixed system of concepts (i.e., word meanings) across changes in theory or belief. So the argument involves an inference from the rejection of analyticity to holism, and from holism to conceptual relativism: theory changes result in a system of thought or perspective that is incommensurable with what it has replaced. Davidson does not seem to take this line of argument very seriously. He writes that “retention of some or all of the old vocabulary in itself provides no basis for judging the new scheme to be the same as, or different from, the old”; the argument we are considering establishes merely the obvious conclusion that “the truth of a sentence is relative to a language” (VICS, 189).

We can of course imagine languages that share many or all of the same terms, but which attach different meanings to those terms—e.g., in some future version of English, “snow” might refer to flowers, “flower” to the number 2, and so on. This is not, presumably, a difference in conceptual scheme, since we could translate that language into our own. But it would seem that this is the kind of difference envisioned by philosophers who argue from holism to conceptual relativism. Davidson concludes that “instead of living in different worlds, Kuhn’s scientists [for example] may, like those who need Webster’s dictionary, be only words apart” (VICS, 189).

This quick response has some force as applied to the case of changes in scientific theory, but fails to address the deeper issues. Suppose that, as Quine thinks, rejection of the analytic/synthetic distinction implies that any and all changes in belief or theory will induce corresponding changes in meaning or conceptual framework. In that case, it is
question begging to suggest that speakers with different theories or beliefs can simply translate each other’s sentences or terms, by creating new dictionaries. If I am to translate or interpret your utterance of a sentence, I must already possess or be able to acquire your concepts: in Davidson’s terms, I must be able to produce a sentence of my own that has the same truth condition as yours.

But the view under discussion is that, if our beliefs are different in any way, we share no concepts. This seems to imply that there is no way for us to entertain precisely the same thoughts, i.e., to utter sentences having exactly the same truth conditions. Your utterance of “Snow is white” is true if and only if snow is white, while mine is true if and only if snow* is* white*. This was the problem Davidson indicated without resolving in asking whether we could really make sense of what the ancients (or some of them) believed about the earth. Their beliefs concerning the earth—or rather, certain beliefs of theirs that we would be tempted to say were about the earth—were inter-related in ways unlike our own, with the result that we have no way of specifying the contents of those beliefs. Or so it would seem, on the assumption of total holism (and the assumption that total holism is implied by the denial of analyticity).

But if this is right, it is simply false that speaker’s with different beliefs or theories need only consult a dictionary to interpret each other’s utterances. In order to create a dictionary of that kind, one would have to trace differences in meaning between two languages or sets of sentences back to some basic, shared system of concepts or meanings. One would have to discover, for example, that the term “earth”, as it figures in modern English, has a certain meaning that is different from that of the corresponding terms in ancient languages. But that could only be discovered by someone who understands both the ancient and the modern concept or meaning. But that knowledge, in turn, would surely require that both of these concepts or meanings can be understood by reference to some more basic system of concepts. And this is precisely what the thesis of total holism says is not possible.

That is, one might, intuitively, think that the difference in meaning could be explained by saying something like the following:

(E) The modern term “earth” expresses the concept of a part of the solar system, a system partly identified by the fact that it is a gaggle of large, cool, solid bodies
circling around a very large, hot star. Certain ancient terms corresponding to the term “earth” expressed the concept of a large flat surface at the center of the cosmos. If one could say something of this kind, one could produce a Modern-Ancient dictionary. But if total holism is true, there are no words in our language, no concepts in our scheme, that capture any of the meanings or concepts of the ancients. We could no more say what the ancients meant by words that we would, intuitively, take to be a translation of “flat”, “surface” or “cosmos” than we could say what they meant by words that we would take to be translations of “earth”. If we have no common concepts or meanings, on the basis of which to explain local differences in meaning, we are more than “words apart” from those whose beliefs are different from ours.

In dismissing the argument from holism, Davidson seems to underestimate the implications of Quine’s attack on the analytic/synthetic distinction. Nevertheless, the line of argument that he dismisses in this way seems to be implicit in his own discussions of meaning holism, which, as we’ve seen, appear to commit him to the extremely radical thesis that any change in belief induces a change in meaning. But there is no real danger of conceptual relativism here. The consequences of total holism are so extremely implausible, or perhaps simply unintelligible, that we should take these to constitute a reductio of total holism.

VII

So far, we have considered Davidson’s treatment of the idea of a partial or local difference in conceptual scheme. On the one hand, it might be that Davidson is concerned here merely with the idea that the concepts associated with a wholly intelligible foreign language could be, as he says, “radically different from our own” (VICS, 195). For reasons already considered, this claim does not seem to be worth arguing about. On the other hand, a more interesting and relevant suggestion would be that some fragment of an otherwise intelligible language might be uninterpretable to us, and Davidson’s discussion could be taken to provide an argument from the holistic nature of meaning and belief for the conclusion that this is not possible. We’ve seen that the
argument from holism is either inconclusive or self-defeating, depending on how holism is understood.

These conclusions are not surprising, since they show merely that we, with our actual capacities and concepts, might be incapable of understanding the beliefs and concepts of others. So long as “we”, whoever that might be, are not assumed to be omniscient, it would be quite natural to grant that this is a logical possibility. We should emphasize, as before, that this is not to grant the possibility of a difference in conceptual scheme in the interesting sense that our incapacity would be explained by purely conceptual facts, i.e., facts about how our beliefs or concepts are related to those we are incapable of understanding. At most, we’ve found that Davidson has not provided a good argument for the conclusion that the nature of our actual concepts, or the mere fact that we have concepts, somehow guarantees the possibility of understanding all others. This would be a profoundly counter-intuitive conclusion, just as counter-intuitive, in my view, as the thesis that this could be somehow impossible simply because we have the concepts and beliefs that we do.

In this section, I will turn from the issue of partial or local differences in conceptual scheme, which we’ve found to be of little importance, ultimately, to a different line of argument from the methodology of interpretation against the very idea of a conceptual scheme. The argument is based on Davidson’s contention that massive error is impossible: any system of belief or conception of the world is by nature largely true. If that is right, then even if others have concepts very different from ours, their system of beliefs, in particular, must nevertheless be largely true. But, for reasons we’ve already considered, the idea that two beliefs (or sentences, or other vehicles of truth) could be incommensurable does not seem to be intelligible.

Seen in this light, we could take the main conclusion of the last chapter to be that if some system of beliefs is true, it is not associated with an alternative to our own conceptual scheme: it is, at most, different and additional to our view of the world. That is not to say, however, that any system of beliefs associated with a purportedly alien conceptual scheme must be true. Some argument is needed for this further conclusion, especially since, as Davidson notes, proponents of the scheme idea are likely to say that truth is “relative” to a conceptual scheme. Or it might be suggested that truth is
immanent, as Quine says at times, that the concept of truth has no intelligible application beyond our language or view of the world. The argument against the possibility of massive error establishes this further conclusion.

So the argument I will explore here can be seen as roughly parallel to Davidson’s argument from Convention T. Both arguments attempt to show that alternative schemes are (largely) true conceptions of the world, and that the concept of truth implies the possibility of understanding across schemes, with the consequence that there is at most one scheme. There is the important difference that the argument from Convention T is unsound, whereas I’ll try to show that the present argument may be sound. At any rate, it rests on more attractive premises. Davidson’s argument against the possibility of massive error begins from his claim that interpretation requires not merely “some foundation” in agreement, but also that “much of what is agreed must be true if some of it is to be false”\(^{163}\) Let us consider this first part of the argument.

Davidson thinks that the far stronger conclusion, that interpretation truth and not just agreement, is also a consequence of “the underlying methodology of interpretation”:

It may seem that the argument so far [i.e., the argument for belief holism] shows only that good interpretation breeds concurrence, while leaving quite open the question whether what is agreed upon is true. And certainly agreement, no matter how widespread, does not guarantee truth. This observation misses the point of the argument: the basic claim is that much community of belief is needed to provide a basis for communication or understanding; the extended claim should then be that objective error can occur only in a setting of largely true belief.\(^{164}\)

Davidson supports this “extended claim” by a thought experiment. We are to imagine an omniscient interpreter, someone who knows everything except what we believe and mean by our words (hence, strictly speaking, a quasi-omniscient interpreter). Such an interpreter must apply the same method as anyone else. To interpret our utterances and discover our beliefs, true or false, he must assume that we are largely in agreement with him. In this case, however, the necessary basis of agreement between interpreter and speaker guarantees that the speaker’s view of the world is largely true.

\(^{163}\) “Method of Truth”, 200.
\(^{164}\) Ibid.
The idea of massive error is unintelligible, then, because “to suppose it intelligible is to suppose there could be an interpreter (the omniscient one) who correctly interpreted someone else as being massively mistaken, and this we have shown to be impossible”.\footnote{\textit{Method of Truth"}, 201.}

I believe that this argument is sound, although it may somewhat misrepresent Davidson’s position, which can be defended without appeal to the idea of an omniscient interpreter. Before trying to clarify and defend that position, I will try to defend Davidson’s argument from the possibility of an omniscient interpreter against some important objections, since it seems to me that the omniscient interpreter argument is at least heuristically useful. I’ll then try to clarify what I take to be the real basis of the argument, which does not concern interpretation so much as the conditions for acquiring language, concepts and beliefs.

VIII

Bruce Vermazen objects that we can envision massive error, omniscience aside, simply by supposing that in addition to the beliefs we ascribe to some speaker and identify as false, many of the beliefs we share with him are \textit{actually} false, unbeknownst to us. Of course, no one can coherently assert that he believes that \textit{s}, and that \textit{s} is false; but presumably an interpreter can imagine that some of his beliefs are in fact false, so long as he doesn’t claim to know which ones. To imagine a case of massive error, then, we need only imagine that the \textit{sum} of false beliefs we ascribe to some speaker, comprising both those that we identify as false and those that are actually false, unbeknownst to us, is greater than the sum of true beliefs. Omniscience adds nothing to Davidson’s argument, if Vermazen is right, since “an omniscient interpreter isn’t barred from imagining that he has false beliefs; I can imagine that I am four feet tall”.\footnote{Bruce Vermazen, “The Intelligibility of Massive Error”, \textit{The Philosophical Quarterly}, 33 (1983): 72.}

Suppose, then, that one can imagine the truth of the following proposition:

(P) The sum of identified and unidentified false beliefs that I have ascribed to some speaker \textit{S} is greater than the sum of true beliefs I have ascribed to \textit{S}.
Does this establish that, *pace* Davidson, massive error is possible? Vermazen’s reasoning seems to be that, since one can imagine the truth of P, it is at least possible that P is true: it is at least possible that the sum of false beliefs ascribed to S is in fact greater than the sum of true beliefs. If so, massive error is at least logically possible.

But does the possibility of imagining the truth of P imply the possibility of P? Certainly it’s doubtful that this is generally true, i.e., that whenever one can imagine that s, it is possible that s. If I make a mistake somewhere in a long series of calculations, it appears that I come to believe that a proposition that is necessarily false. To believe it, presumably, implies at least the ability to *imagine* or *conceive* of its truth, in some sense, even though it isn’t possible that the proposition in question is true. So it may be that the truth of P is imaginable, in some sense, even though massive error is impossible. Perhaps, as with beliefs formed by a mistake in calculation, it isn’t even *logically* possible that P is true, and yet it is possible to imagine that it is.\(^{167}\)

However, it might be argued that if some proposition is necessarily false, it couldn’t be truly intelligible. It might appear to be intelligible, but if it is necessarily false, it must turn out, on closer inspection, to be unintelligible. On this view, then, something that seems intelligible often is not, in fact, intelligible; but then true intelligibility is as difficult to establish as logical possibility. It is sometimes obvious that some proposition is or isn’t intelligible, is or isn’t logically possible. The proposition that Bruce Vermazen is not Bruce Vermazen is obviously unintelligible if any proposition is. (Or, if propositions must be intelligible, at least in principle, we should say instead that the sentence “Bruce Vermazen is not Bruce Vermazen” is obviously unintelligible.) Other cases are less clear, however—e.g., the proposition that Bruce Vermazen’s parents were not his actual parents. It isn’t obvious that this could be true, and therefore it isn’t obvious that it is intelligible, if we take intelligibility to require possibility. But P is not *obviously* intelligible (or logically possible). So, for all we know, it may appear to be

\(^{167}\) If this is true, then Davidson is wrong, strictly speaking, to say that the idea of massive error is unintelligible. But if this is something like the intelligibility of the proposition that \(2 + 2 = 5\), whatever exactly that kind of intelligibility may be, his basic position is unaffected. The main point is that massive error is impossible—or, alternatively, it is that the idea of massive error is not *ultimately* intelligible, on careful reflection. Neither precludes the possibility that, as with other ultimately unintelligible ideas, the idea of massive error has some kind of *prima facie* intelligibility, whatever exactly that may come to, on reflection.
intelligible despite being unintelligible, and in that case, it would not follow that it is logically possible that P is true.

If we assume that the intelligibility of some proposition implies that the logical possibility that that proposition is true, we should allow that our intuitions or initial conclusions as to intelligibility are often wrong. And P seems to be the kind of proposition that we might well be wrong about. So, while Vermazen’s inference from the intelligibility of P to the possibility that P is true is valid on this reading, the same reading leaves us with no good reason to suppose that P is intelligible. But if we assume, instead, that our intuitions as to intelligibility are always or usually reliable, we have to concede that some necessarily false propositions are apparently intelligible, and in that case, Vermazen’s inference is invalid: the fact that an interpreter can imagine that P is true does not imply that it is logically possible that P is true.

The question of whether the apparent intelligibility of P is real intelligibility brings out a more basic flaw in Vermazen’s argument. Clearly we can, as he suggests, “look back at the vast changes in common beliefs through history and reflect that further changes are in prospect”. In this way, we can probably make sense of the skeptical possibility that many of our current beliefs are actually false, unbeknownst to us. But how can we know that this kind of skeptical reflection enables us to imagine that enough of our beliefs are actually false so as to yield, when added to those that we identify as false and ascribe to someone else, a massively wrong view of the world? By hypothesis, we can’t count them, or in any other way measure the ratio of actual truth to falsity in our current system of beliefs. For any measure of that ratio would require identifying particular beliefs that we presently hold as false beliefs, which we can’t do, on pain of incoherence.

However, if we can’t make any kind of estimate of the extent to which our current system of beliefs is actually false, neither can we know whether we are actually imagining that P is true. To imagine that P is true would be, again, to imagine that the sum of beliefs we ascribe to some speaker S is, on balance, more false than true. But that sum consists of two others: the sum of identified false beliefs and the sum of unidentified false beliefs. Since we have no way of discovering the magnitude of the second sum—

---

168 “Intelligibility”, 72.
again, regardless of how “quantities” of belief might be measured—we have no way of knowing the magnitude of the second sum when added to the first. But if we can’t know that, we can’t know that this latter magnitude is greater than that of the beliefs ascribed to the speaker that we take to be true. All we are able to imagine, therefore, is that there is some degree of falsity, greater than the degree of falsity that we can measure, in the system of beliefs we ascribe to S. But, for all we know, the degree of falsity under discussion is actually far less than what would be required for massive error.

The skeptical interpreter’s position could be either of these:

(1) The interpreter ascribes to S a massively false system of beliefs, and forms the true belief that the system of beliefs ascribed to S is massively false.

(2) The interpreter ascribes to S a largely true system of beliefs, and forms the false belief that the system of beliefs ascribed to S is massively false.

To imagine that P is true would, of course, be to imagine that (1) is true. But unless the interpreter can know the system of beliefs he ascribes to S is massively false, he can’t know whether he is, instead, imagining that (2) is true. And he can’t know this, since this would require him to believe and disbelieve the same propositions (or hold true and false the same sentences). For all he knows, then, the interpreter is imagining that (2) rather than (1) is true. But then, even assuming that conceivability or intelligibility generally implies possibility, Vermazen’s argument is invalid. The interpreter’s ability to imagine that (2) is true would then entail the possibility that (2) is true, but this would not establish that belief in massive error is intelligible (and therefore, that massive error is possible).

In order for that latter conclusion to follow, we would need the principle that if it is intelligible (conceivable, imaginable) that someone forms the belief that s, the belief that s is intelligible. And given this principle, together with the premise that the interpreter can make intelligible (2), or imagine that (2) is true, it would follow that it is intelligible that someone is in massive error, and hence that it is possible that someone is in massive error. However, this principle seems false. I can imagine acquiring the belief that God is able to make a square circle, but this doesn’t require that I consider that belief itself to be intelligible, let alone that it is logically possible that the proposition believed is true. Surely it doesn’t follow that what I can imagine believing is itself intelligible.
IX

I now turn to a second important objection to Davidson’s argument from the possibility of an omniscient interpreter, according to which the argument begs the question in presupposing that an omniscient interpreter would be able to interpret our words. Kirk Ludwig suggests that the conclusion of Davidson’s argument should be merely that “either we are mostly right and correctly interpretable by the omniscient interpreter, or we are largely wrong and so not correctly interpretable by the omniscient interpreter”.\textsuperscript{169} If we accept Davidson’s first premise, interpretation requires substantial agreement. But then the assumption that an omniscient interpreter could succeed in making sense of our utterances and other behavior is acceptable only on the assumption that we are largely right about the world. So it seems that we can’t have any reason to accept this premise of Davidson’s argument unless we already have some reason to accept its conclusion.

I think Ludwig is right about this. But how can the issue be framed without begging the question? Suppose that we want to leave it undecided whether the omniscient interpreter would be able to make sense of our utterances and behavior. In that case, it would seem that we are allowing the possibility that we might be massively wrong—entirely wrong, even. And yet this possibility is hard to distinguish from the possibility that our words are simply meaningless, that we believe nothing at all. For the idea is that the interpreter, knowing every non-mental truth about the universe, cannot discern any systematic relations between our words and anything in the extra-mental world, regardless of how carefully and extensively he studies our behavior. Then there are no such relations: no objective situations in which our sentences are true or false, no objects of which our singular terms and predicates are true.

If this were in fact our situation, it would be hard to make sense of the initial assumption that we have beliefs at all. At least, this is what I will argue. Belief would seem to involve, by its very nature, some kind of normative constraint imposed by the

world. This is true of false belief just as much as true belief. If I mistakenly believe that zebras are cows, I should be disposed to say, “There’s a cow” in the presence of zebras. If I mistakenly believe that water is poisonous, I should be averse to touching or drinking water, and so on. By contrast, if there are no such correlations between behavior and dispositions, on the one hand, and the changing conditions of the world, on the other, we have no clear sense of what it is for me to have any view of the world at all, true or false. But this is what is required by the assumption that an omniscient interpreter could not discover our beliefs.

Ludwig’s conclusion is therefore not quite right. The intelligible alternative to the possibility that “we are mostly right and correctly interpretable by the omniscient interpreter” is not that “we are largely wrong and so not correctly interpretable by the omniscient interpreter”. It is that we are neither right nor wrong, to any degree, because we do not have beliefs and our utterances are strictly meaningless. If so, we are not interpretable, correctly or incorrectly, by the omniscient interpreter, because there is nothing to interpret. One might conclude from this characterization of Davidson’s argument that it allows for an even more radical form of skepticism than the traditional, Cartesian skepticism it is supposed to counter.

But it is not debatable, surely, that we at least appear to have beliefs, desires and intentions, that our words appear to have some meaning. Nor does this beg the question against the proponent of this more radical form of skepticism, since we are conceding, for the sake of argument, that this appearance might be some kind of illusion. But it would be incoherent to hold that mindless creatures mistakenly believe that they have minds—that, for example, it appears to us that it appears to us that we have minds, when in fact we are mindless and the world does not so much as appear to us to be a certain way.

Ironically, then, the best defense of Davidson’s anti-Cartesian position, then, may ultimately rest on something like Descartes’ cogito. No one could possibly be wrong if he thinks that he is presently thinking, believing or perceiving something, or meaning something by his words, for merely to believe this guarantees the truth of the belief. This would seem to guarantee that we do in fact have beliefs, and that we would be (correctly) interpretable by the omniscient interpreter, and that, therefore, our beliefs must be largely true. Davidson’s presentation of his argument against the intelligibility of massive error
is thus rather misleading. The basis of the argument is not that interpretation requires agreement, but that beliefs depend for their identity and content on the extra-mental world.

As Davidson puts it in “Mental Events”, “global confusion, like universal mistake, is unthinkable, not because imagination boggles, but because too much confusion leaves nothing to be confused about”\textsuperscript{170} That error is possible only against a background of true belief is not, then, an “extended claim”\textsuperscript{171} drawn from the premise that interpretation yields agreement. Rather, that interpretation yields agreement and that massive error is unintelligible are both consequences of the dependence of meaning and belief on the objective world.

X

On the present reading of Davidson’s argument, the argument depends on a key premise that is not obvious in his discussion of the omniscient interpreter; this a premise to the effect that meaning and belief are in some basic sense constituted by the believer’s objective circumstances. Davidson often presents his argument in a way that suggests that this externalist conception of meaning and belief is a consequence of the methodology of interpretation: because the interpreter must take the speaker’s beliefs to be correlated with what he takes to be the objective world, the speaker’s beliefs must actually be so correlated. My suggestion is that we view this latter claim about the methodology of interpretation as a consequence of Davidson’s externalism. Or, to put the point more cautiously, we could view both of these as mutually reinforcing positions that, taken together with Davidson’s other views, are supposed to provide an elucidation of the key concepts of meaning, belief and truth.

Why exactly must beliefs be tied somehow to the objective world, in a way that ensures that they are largely true? The answer is to be found in Davidson’s account of what he calls “triangulation”, which he takes to be a necessary condition for meaning and belief, and so for interpretation. The meaning of an utterance (and, by extension, the

\textsuperscript{171} “Method of Truth”, 200.
belief expressed by the utterance) is determined by the *actual causes* of utterances and beliefs. In particular, by whatever it is that systematically causes both speaker and interpreter to form a certain belief or be disposed to utter certain words. Causality is what provides for the possibility of a determinate content, of thought and language being about something in particular, while the requirement that the cause be correlated with the thoughts and reactions of others is supposed to fix *which* determinate content a thought or utterance has.

As in the last chapter, Davidson’s use of the term “interpretation” here tends to obscure the real issues. For interpretation, presumably, is possible only for two people, a speaker and an interpreter, who already have beliefs, concepts and language. Thus, one might object that Davidson’s appeal to causality begs the question, in much the same way that the original argument from the methodology of interpretation appears to do. That is, the original argument seemed to involve an inference from the premise that interpreter and speaker must agree to the conclusion that what they agree must be true; but this seems invalid.

Thus, Davidson at times seems to beg the question against the skeptic by assuming that what an interpreter and a speaker *take* to be the causes of their reactions and thoughts must be the actual causes:

What stands in the way of global skepticism of the senses is, in my view, the fact that we must, in the plainest and methodologically most basic cases, take the objects of a belief to be the causes of that belief. And what we, as interpreters, must take them to be is what they in fact are.\(^{172}\)

But the fact that we do or must take the cause of an utterance to be, say, a rabbit’s having appeared in front of the speaker, does not imply that that is in fact its cause. Two speakers with similar delusions might believe each other’s utterances to be caused by imaginary rabbits. Unless we already know that any speaker’s beliefs must be largely true—in particular, beliefs about the causes of other people’s utterances and reactions—we have no reason to accept Davidson’s conclusion.

---

Barry Stroud has emphasized this line of objection. As he sees it, Davidson’s argument against massive error establishes at most that belief ascription is largely truth-ascribing; the additional appeal to causality adds nothing to the original argument from the methodology of interpretation. He focuses on the fact that talk of interpretation and communication is intelligible only if we assume that the speakers in question already have beliefs with particular contents. Davidson’s position would seem to be that both the interpreter and the speaker must both have certain similar beliefs about the causes of each other’s utterances, and their own, if the contents of other beliefs are to be determined by triangulation. This provides more detail about what is required by interpretation, but no new reason to think that what interpreters and speakers (must) believe is true.

Although he does not put his point in precisely these terms, we could express Stroud’s thought by saying that the notion of causation that figures in this account is not really the notion of causality per se, but the notion of believed or ascribed causality, which does nothing to advance the argument:

The notion of cause plays its role in interpretation within the scope of the psychological verbs expressing the propositional attitudes of interpreters and speakers. A interprets B by having beliefs about what causes B’s utterances, and B does the same for A… It does not follow from “A believes that B’s utterance is caused by the presence of a rabbit” and “B believes that A’s utterance is caused by the presence of a rabbit” that their utterances are caused by the presence of a rabbit.173 Just as truth can be ascribed to false beliefs, the property of having caused an utterance can be ascribed to things that are not its causes. Unless, of course, we are to presuppose that in such cases what is believed must be true. Thus, we seem to be moving in a circle.

I think this reading of Davidson’s argument is wrong, although it is understandable given his misleading use of the term “interpretation”. The following passage from Davidson’s reply to Stroud illustrates both what is misleading in his presentation of his argument and, I think, a more interesting and powerful reading of the argument that does not beg the question:

What is essential to the ostensive process is that when the interpreter is caused to think there is a pencil, the informant is caused to say “Woozle” [for example]… How does the skeptic view this situation? Can both interpreter and informant be mistaken regularly? From the start? No; as Wittgenstein points out, the interpreter (learner) is not in a position to doubt the ostensions he observes. For him, the new sentence is being given a content… There is no crack here where the sceptic can drive a wedge between belief and reality… It is only after a belief has a content that it can be doubted. Only in the context of a system tied to the world can doubt be formulated.174

If we take Davidson’s remarks at face value, they are inconclusive at best, and quite confusing.

The “start” to which Davidson refers is the stage at which a speaker first comes to have a language and a system of beliefs and concepts, on the basis of no prior language or beliefs. This is quite clearly indicated by Davidson’s earlier remark that although there could be, as Stroud suggests, speakers who communicate on the basis of shared false beliefs—e.g., a shared false belief that a rabbit is in view—this could not be an explanation of how they came to have beliefs in the first place:

How did the discussants come to have the concept of a rabbit? Or, if they learned this concept indirectly, by being told what a rabbit is, or seeing a picture, then how did they acquire the concepts on which these learning procedures depend? At some point they learned enough concepts through ostensive occasions to anchor their sentences to the world, otherwise there would be no way to account for the contents of their perceptual beliefs and the sentences that express them.175

So we are concerned here, ultimately, with the kind of situation in which beliefs and concepts are first acquired, and not with the situation in which, having certain beliefs and concepts, one attempts to discover what someone else believes or means by their words.

It is absurd to speak of “the interpreter (learner)” [my emphasis] in such a situation. A child learning a first language is not learning what expressions in the language mean by translating or interpreting them, i.e., understanding these by means of

---

174 Donald Davidson, “Reply to Stroud” in The Philosophy of Donald Davidson, 165.
175 Ibid., 164.
some other language or system of concepts and beliefs. There may be a learner in this situation, but not an interpreter. This difference is crucial to the point I want to make on Davidson’s behalf. It is precisely because a pre-linguistic child is not an interpreter of other people’s sentences that “there is no crack here where the skeptic can drive a wedge between belief and reality”.

As Davidson emphasizes, what is happening here is rather that a child’s mental states and verbal behavior are being given content. Until they have content, they can’t be doubted; nor can they be coherently described as wrong or false. The child does not hypothesize that terms have meanings, or believe that sentences are true under certain conditions. Not yet having concepts, he can’t, of course, frame theories about which concepts are associated with which words, etc. So he can’t falsely believe that words refer to something other than whatever it is that, together with the teacher’s reactions and encouragement, causes his utterances. If the learner is caused to say “Woozle” when pencils are in view, and not otherwise, pencils must be the actual cause of his utterances of “Woozle”.

We could say, then, that “what stands in the way of global skepticism of the senses” is “the fact that we must, in the plainest and methodologically most basic cases, take the objects of a belief to be the causes of that belief”, if the plainest and most basic cases are taken to be, not cases of radical interpretation, but cases of first language acquisition. The teacher may have false beliefs about the causes of the learner’s utterances, from time to time. But if he is never right about those causes, and never able to develop a largely true and systematic understanding of how they are related to the learner’s reactions, he will never be able to teach the learner the use of any words or sentences. If he does succeed, his beliefs about the causes of the learner’s utterances must have been often true, and systematically related to one another. It is in this sense that what the teacher takes to be the causes must be the causes: his judgement is authoritative, because he is fixing aspects of the environment that the learner’s beliefs will have as their contents.

The contents acquired in this way are not mental entities in the traditional sense; they are not in the mind of the learner, in some sense that contrasts with what is in the learner’s non-mental environment. Rather, the contents of the learner’s initial utterances
are states of the environment, or objects in the environment. The learner associates
sentences with their truth conditions, which are the contents of those basic sentences.
This is presumably the meaning of Davidson’s claim that in the basic cases of
triangulation, speakers must acquire “concepts through ostensive occasions” that “anchor
their sentences to the world”. Certain sentences in anyone’s vocabulary will be
anchored to the world in that their contents are, at least in part, things or states of things
in the world.

This is not intended as an account of sufficient conditions for the emergence of
language and thought. The point is simply that if all the contents of our beliefs were
entirely internal to the mind, as some philosophers have believed, there would be no way
of explaining how we acquired those contents, or why they should be taken to be about
the extra-mental world at all. But if this kind of externalism is correct, it does not seem
intelligible to suppose that we could be massively mistaken about the world, or that we
could have incommensurably different ways of interpreting or representing the world.

As in the last chapter, then, we see that Davidson’s appeal to an undifferentiated
notion of interpretation is confused. Just as we must have a non-interpretive
understanding of our own utterances and mental states, so the process by means of which
we become speakers and thinkers in the first place must also be non-interpretive. The
view defended here complements what the view defended earlier concerning non-
interpretive knowledge of what one’s own utterances means. I argued that such
knowledge consists in knowledge that a sentence is true under a certain (possible)
condition of the world, and not in the interpretive knowledge that it has the same truth
condition as another sentence. The present suggestion is that this knowledge, in the basic
cases, associates sentences with the actual conditions under which they are true: “in
giving up the dualism of scheme and world, we do not give up the world, but re-establish
unmediated touch with the familiar objects whose antics make our sentences and opinions
true or false” (VICS, 198).

---

176 “Reply”, 164.
CONCLUSIONS

In this essay, I’ve explored the idea of a conceptual scheme by way of an evaluation of Davidson’s influential discussion, and defended a version of his thesis that the idea of a conceptual scheme is not intelligible. While many of the arguments I’ve set out are inspired by Davidson’s discussion, my position differs in some important ways from his. My view does not rule out the possibility of a language that we would be unable to interpret, which I think is in fact an intelligible possibility. But I don’t think that this is equivalent to the claim that there could be an alternative conceptual scheme. Our inability to interpret a language would imply a difference in scheme only if our inability were due to the relation between our language and the other (i.e., the relation of “incommensurability”).

Other important differences concern the way in which I think the thesis that there could not be alternative schemes should be supported. I’ve rejected a number of the subordinate arguments that Davidson sets out against the intelligibility of a difference in conceptual scheme. Finally, I’ve also raised some objections to important features of Davidson’s broader philosophy, which are often taken to figure in his attack on the scheme idea: radical meaning holism, and the assimilation of understanding to interpretation. I find these positions unconvincing, and in any case I don’t think that they contribute to the case against the intelligibility of the scheme idea.

Despite these many differences, however, I have accepted Davidson’s general strategy against the intelligibility of the scheme idea, which I take to be a development of his initial observation that conceptual relativism involves a paradoxical pair of claims. Supposedly, divergent schemes are directed at a single, common entity, and yet this common entity radically under-determines the character of the schemes that serve to interpret or represent it, with the result that interpretation is impossible. The argument is intended to show that, as applied to language or conceptual thought, this picture is not just paradoxical but incoherent.
Conceptual schemes are to be identified with sets of (inter-interpretable) languages, Davidson thinks. So two schemes differ, if they do, only if interpretation is impossible across the languages with which are associated. But there is no way to make sense of the idea that interpretation is impossible (in principle) across certain languages. All languages are interpretable by means of our language, and hence interpretation across any two languages is possible, at least in principle. So there is at most one conceptual scheme. But the very idea of a conceptual scheme is empty unless we have some notion of an alternative. Therefore, talk of conceptual schemes has no real content or purpose. I think that this general line of argument is sound, despite my misgivings about Davidson’s subordinate argumentation, and about his own understanding of what the argument establishes.

In this chapter, I’ll review Davidson’s treatment of the scheme idea and what I think are the main critical points raised in my evaluation of his views. Then I’ll discuss in more detail the different position that I’ve defended concerning the idea of a conceptual scheme, and consider how my conclusions bear on the larger philosophical issues that Davidson takes to be closely related to the scheme/content distinction.

I

In denying that languages embody or encode “conceptual schemes”, Davidson means that languages do not somehow shape or filter our experience of the world. He holds, instead, that the subjective and the objective cannot be prized apart, that no sharp or fundamental distinction can be drawn between the contents of the mind and the extra-mental world about which we take ourselves to be thinking and communicating. For, as we’ve seen, he argues that those contents are to be identified with objects and events in our environment (in the epistemologically and semantically basic cases, at least). On this view, then, the shared environment required by the scheme idea guarantees that any differences between thinkers can be construed as differences of meaning or belief arising within a single, shared system of concepts: they can always be described and understood by means of a single language.
Hence Davidson’s initial observation that “the dominant metaphor of conceptual relativism, that of differing points of view, seems to betray an underlying paradox” (VICS, 184). According to the conceptual relativist, thinkers with different conceptual schemes inhabit a single world, but conceive of it in incommensurable ways. For Davidson, though, the mere fact that they inhabit a single world means that their mental lives must be constituted, at least in part, by that same world, and this would seem to imply the possibility of translation. Davidson then tries to flesh out this initial objection to the intelligibility of the scheme idea by exploring in more detail the ways in which conceptual schemes are said to differ.

As we noted earlier, this is also an exploration of the real features of language and thought that figure in various accounts of conceptual schemes, such as reference and truth. So Davidson aims to show that, as he says, “different points of view make sense, but only if there is a common co-ordinate system on which to plot them” (VICS, 184). Language or conceptual thought can indeed be taken to involve a point of view, and different languages or systems of concepts can be likened to differences in point of view. The various metaphors of conceptual relativism suggest how we might make this thought more precise: that is, how we can “say what the relation is” between scheme and content, and “be clearer about the entities related” (VICS, 191). But when we consider more closely what these putative schemes and content might be, and how they might be related, we will also find that the entities and relations under discussion guarantee the possibility of interpretation, by reference to the shared environment within which purportedly different conceptual schemes must operate.

Davidson frames the debate in terms of two dualisms, which he takes to be suggested by talk of conceptual schemes: the dualism of the analytic and the synthetic, and the dualism of conceptual scheme and empirical content. The notion of analytic truth suggests “a distinction within language of concept and content” (VICS, 187), which can be retained even if the analytic/synthetic distinction is rejected. For language as a whole may then be contrasted with empirical content: experience, stimulation, the world, the facts, etc. Language is seen as an “organizing system”, the empirical content as “something waiting to be organized” (VICS, 189). Davidson writes that this “dogma of empiricism … cannot be made intelligible and defensible” (VICS, 189).
We’ve seen that this way of characterizing the scheme/content dualism is misleading, since Davidson associates the dualism with positions that have no obvious connection to empiricism—most importantly, with radical skepticism, representationalist theories of mind and language, and the correspondence of truth. In any case, the basic idea is that if it makes sense to speak of a conceptual scheme, others must be at least possible. Assuming that there is an essential link between conceptual thought and language, we can then formulate two necessary conditions for a difference in scheme. First, both languages or schemes must be related in some specified manner to “something neutral and common that lies outside all schemes” (VICS, 190), and, second, interpretation must be impossible across these languages.

Davidson then pursues the question of whether this is an intelligible scenario by considering in more detail the metaphors of conceptual relativism (the subordinate metaphors, so to speak). To summarize, if we are to make sense of the scheme idea, we need to make sense of the idea of differences in scheme. Having set out an abstract account of the conditions for such a difference, Davidson then examines various accounts of the scheme/content relation that are supposed to explain how a situation satisfying both conditions could arise. The conclusion is that languages characterized in this way are interpretable or translatable by means of our own language.

II

In this section, I’ll review the main points of my evaluation of Davidson’s arguments against various particular accounts of the scheme/content relation. In the next, I discuss some more general arguments against conceptual relativism that can be drawn from his broader philosophy. The arguments in the first group are based on Davidson’s view that there are basically four ways in which the scheme/content relation can be characterized. If languages are conceptual schemes, then they either organize or fit some common entity, which may be either experience or reality. Davidson explores the relations between language and experience or reality suggested by these different metaphors, and argues that, insofar as we can make sense of these relations, interpretation across putatively different schemes will turn out to be possible.
The metaphor of organization seems to concern reference (broadly construed). If one language contains predicates applicable to some range of entities, or singular terms referring to certain entities, it is logically absurd to suppose that some other language related in the same way to that ontology is incommensurable with the first. For it must be possible, in this kind of situation, to frame the differences between the first two languages by means of a third language. In this way, supposedly different conceptual schemes can be seen as fragments of a single language appropriate to different aspects of a single range of objects (or perhaps to different objects). The common ontology thus seems to ensure the possibility of interpretation by reference to a common co-ordinate system.

The metaphor of organization might also be understood, though, as a gesture towards the intensional features of language. Although Davidson does not address this construal of the metaphor of organization, it seems likely that this is very much what proponents of the scheme idea intend. The idea need not be that terms in different languages refer to different things or collections of things, but rather that the ways in which they refer to those things are very different. Against this suggestion, I’ve argued that ways of referring to objects, i.e., senses or concepts, cannot come into any kind of conflict unless they are identified or associated with conditions for the application of terms. This leads to conclusions that preclude a difference in conceptual scheme.

Suppose that an object o is conceived of by one speaker as an F, and by another speaker as a G. If this to be an example of incommensurability, these two thoughts or ways of representing o must be not merely different, but also in some kind of conflict or tension. Only in that case would it be tenable, perhaps, to hold that these different ways of thought are incommensurable, i.e., that merely to conceive of o as an F would make it impossible to conceive of it as a G. But it is hard to see how thoughts could be in conflict or tension except, ultimately, by being somehow inconsistent. Where the principles associated with two concepts applied to the same object are inconsistent, there are only two possibilities.

If there are analytic truths, i.e., truths constitutive of concepts, then at most one of these sets of principles is in fact applicable to that object—the other concept is incoherent and applies to nothing. If there are no analytic truths, the conflict between principles is simply a disagreement, a difference of opinion. But neither of these possibilities is
consistent with the scheme idea. On the first horn of the dilemma, the supposed inhabitants of an alternative scheme are simply confused inhabitants of our own. On the second, they merely have different beliefs about the world, which may complement our own or put them in disagreement with us. This is again an instance of Davidson’s general strategy against the idea of a difference in scheme: the metaphor of different points of view turns out to imply a common co-ordinate system, and hence the possibility of interpretation.

The argument against the idea that languages qua conceptual schemes fit some entity is more complex, as we’ve seen, and depends on a series of conceptual assimilations. For a sentence to fit the evidence, Davidson thinks, is for it to be warranted by the (totality of) evidence or experience, which is just for that sentence to be made true by that evidence, or by the facts, which is simply for it to be true. I’ve argued that these conceptual assimilations are not defensible. The concept of fitting the evidence is an epistemological concept, while the concept of fitting the facts or the world is a metaphysical concept. One might hold that some theory fits the sensory evidence, even the totality of sensory evidence, and yet is false (although, presumably, there could be no reason to believe this). The assimilation of the first notion to that of truth is neither acceptable in its own right nor something that Davidson’s opponents accept (or that they should accept, given their other commitments). So this phase of the argument fails.

However, the present account of conceptual schemes can be dismissed for a different reason that Davidson sets out elsewhere: there is no such thing as non-propositional “evidence”. The evidence that supports a theory must itself be propositional if it is to stand in rational relations to sentences, beliefs, or other vehicles of truth. Mere sensory stimulation may have a causal relation to theories, beliefs and the like, but it cannot warrant or justify acceptance of sentences or theories. If divergent schemes, on the present account, are theories that warranted by, the same range of sensory evidence, they must be inter-interpretable. For if both theories, i.e., groups of sentences, can be inferred from some other group of sentences that states or constitutes the shared evidence, both theories stand in rational relations to that latter group of sentences. But then they also stand in rational relations to one another; they are consistent or inconsistent, and are just fragments of a single scheme.
Despite the failure of the first phase of the argument against accounts of the scheme/content relation based on the metaphor of fit, the remaining phases can be treated as independent arguments against the intelligibility of the scheme idea. One is an argument against the idea that conceptual schemes are (or are associated with) true sentences that fit or correspond to a single reality, world or set of facts—Davidson takes this to be a formulation of the scheme idea. Against this formulation, Davidson deploys the slingshot argument, which is intended to show that there is no way to individuate facts or other entities to which true sentences might be said to correspond. If so, the idea of correspondence to fact reduces to the idea of truth.

While I think that the slingshot does generate serious problems for correspondence theories, I am not convinced that there is any need for this line of argument. As I see it, there is a simpler argument against this formulation of the scheme idea, requiring no controversial or technical assumptions about semantics. Instead, the argument depends simply on an intuitive conception of facts. If there are such things, they must have the same contents as the sentences they make true. It is this that allows facts to play the role of truth makers, and which gives content to the idea of a correspondence: the correspondence consists in identity of propositional content. But then two theories or groups of sentences that correspond to the same facts must be inter-translatable in the strongest sense, since they have the same contents.

III

The last phase of the argument against accounts of scheme/content dualism based in the metaphor of fit concerns the relation between the concepts of truth and translation. Presumably alternative conceptual schemes, if such there be, are associated with certain (incommensurable) truths. Davidson asks whether we can make sense of the idea of a true sentence or group of sentences that is not translatable into our language. We cannot, he thinks, because Convention T embodies some central aspect of the concept of truth, as applied to language. In conceiving of a sentence as true, we are committed to the idea of a truth theory for the language to which the sentence belongs; but the truth theory must also be a translation, in some sense, of that language. This argument can thus be
detached from the larger argument in which it appears as a subordinate argument, and treated as an argument against a more general common feature of all accounts of the scheme/content dualism.

Given the way that Davidson states the argument, it seems to be invalid. Even if we accept Davidson’s claim that ascriptions of truth imply the possibility of translation into some meta-language or other, there is no reason to suppose that this meta-language must be English, or any language translatable into English. Perhaps there could be a language in which truth conditions for an alien conceptual scheme (i.e., group of sentences) can be stated, but that language, like the one it translates, is not translatable into ours. So even if the concept of a true sentence or theory implies the possibility of translation by way of a Tarskian truth theory, it does not follow that the truth theory is one that we could understand or translate into our language.

I’ve suggested a reading of Davidson’s argument that is not open to this objection. On this reading, Davidson’s position is that there could be a meta-language by means of which our language and any other could be treated as fragments a single object language, and hence translated into that meta-language. This seems right, because Convention T does not require translation in any robust sense; mere disquotation is enough. Thus, for any two languages, it is trivially true that there is a meta-language that “translates” both, i.e., states the truth conditions of sentences in both languages.

But this reading of the argument invites a further objection: this kind of “translation” is not enough to ensure that the meta-language in question is one that anyone could understand. For, if we are not simply presupposing that there are no incommensurable languages, the meta-language might consist of incommensurable fragments. I consider the possibility that Davidson would not be troubled by this objection because it would seem that, in his view, all linguistic understanding ultimately reduces to disquotation (or, more precisely, to something that is characterized by disquotation). So, on this kind of view, incommensurability is an empty or unintelligible notion because translation or interpretation by way of a disquotational truth theory is a paradigm of linguistic understanding.

This counter-intuitive view has its sources in Davidson’s philosophy of linguistic meaning, which I argue is not tenable in this respect. My contention is that the very idea
of interpretation requires that there be some form of understanding that is not interpretive, whatever exactly that kind of understanding may be. If this is right, the argument that I have speculated may lie in the background of Davidson’s discussion of Convention T is unsound. The argument depends on the assumption that all understanding is interpretive—which appears to mean, surprisingly, that it rests ultimately on knowledge of disquotational truth theories, hence on a form of translation that would not normally be considered interpretive. I should perhaps stress that it is not at all clear whether this is in fact the kind of reasoning that supports Davidson’s conclusions regarding the idea of a true but untranslatable theory.

As before, I’ve tried to provide an argument for Davidson’s conclusion that rests on less controversial premises. The suggestion under discussion was simply that divergent conceptual schemes might be true but not inter-translatable. In the absence of any indication of how exactly the scheme/content relation is to be understood, it is hard to know what to make of this suggestion. The onus is surely on proponents of the scheme idea to show that this makes sense. Given that the issue here is in fact the possibility of understanding or interpreting across languages, the present suggestion would seem to be that there could be sets of true sentences such that no thinker could understand them all, or make sense of the claim that they are jointly true.

Or, to put the point another way, it would seem that what we are supposed to imagine is this. There are two groups of true sentences, and the meanings or truth conditions of these sentences concern the very same world—presumably they must concern different objects, or different regions or aspects or properties within the same world if they are “incommensurable”. But to say that they are incommensurable is to say, in effect, that although these two groups of sentences are true because of how things are in that single world, there is no way that they could be understood as complementary parts of a single, more comprehensive description or account of the world. (For, again, the idea of incommensurability is supposed to involve a relation between concepts or languages, not between concepts or languages, on the one hand, and the capacities of some particular thinker or observer, on the other.)

But what could it mean to say that it is impossible for any thinker to understand two groups of sentences or theories as complementary descriptions of a single world, if
not simply that those two groups of sentences or theories simply are not complementary? For example, no thinker, presumably, could construe the sentences “Snow is white” and “It is not the case that snow is white” as complementary descriptions of snow, because they are inconsistent. The challenge for this line of thought is then to find some middle ground between (i) the claim that putatively divergent schemes are inconsistent accounts of a single world, in which case they collapse into one; and (ii) the claim that these schemes are consistent and complementary, but not jointly intelligible because of some contingent, non-conceptual reason stemming from the limitations of a particular kind of thinker or observer, in which case they will again collapse into one scheme.

Davidson’s argument against the intelligibility of a partial difference in conceptual scheme introduces new considerations. The argument is not stated very clearly, but seems to depend on meaning holism and the misleadingly named principle of “charity”. Davidson suggests that there could not be a partially uninterpretable language because an interpretation of any part of a speaker’s language requires us, as interpreters, to ascribe to the speaker a system of beliefs and concepts, and to maximize the overall truth and rationality of the system. Any supposed (partial) differences in scheme could then be seen either as straightforward disagreements or, alternatively, as cases in which the alien speakers are using words in unusual ways.

It may be, although it is not entirely clear, that the argument equivocates over different versions of the holism thesis. Belief and meaning might be claimed to be holistic in that to have any one belief is to have a system of belief that links the first belief to countless others. Or, instead, the holistic thesis might be that to have any one belief is to have a system of belief that links the first belief to all other beliefs that one holds (or should hold, or could be interpreted as holding). The argument against the possibility of partial differences in scheme is sound only on the second version of the holistic thesis, i.e., only on the assumption of total holism. But this radical form of holism is extremely implausible, as we’ve seen. Moreover, total holism in fact implies the possibility of a difference in conceptual scheme: if there is any difference at all between the beliefs of two speakers, on this kind of view, neither will be able to understand anything that the other says.
However, a more moderate form of holism leaves open the possibility of a partially uninterpretable language. For even if we suppose that to ascribe a single belief to a speaker requires the ascription of an entire pattern of belief within which that one belief can be identified and given content, this does not imply that the pattern comprises all of the speaker’s beliefs—or all the beliefs he could be interpreted to have, or all the interpreter’s beliefs, etc. Even if the pattern ascribed in the first case comprises an infinite number of beliefs, which is presumably never the case, there is no reason why there could not be other patterns conceptually isolated from that one. (That is, the claim that a certain group of beliefs is infinite in number does not imply that all beliefs, of the speaker or the interpreter, are in that group.)

In light of considerations initially introduced by this phase of Davidson’s argument, concerning the nature and methodology of interpretation, I’ve discussed how Davidson’s form of content externalism can be deployed against the idea of a difference in scheme. For Davidson, language learning and communication depend ultimately on our grasp of truth conditions supplied by an inter-subjective situation in which at least two speakers are caused to respond verbally to a shared stimulus. I’ve considered this more general line of argument by reference to Davidson’s argument against the possibility of massive error, which seems to rest on two basic normative constraints constitutive of the concept of belief.

First, there is the holistic constraint that beliefs are by nature rationally related to other beliefs, which takes to be constitutive of belief. But as a matter of logic, at least, a rationally coherent system of belief could be largely or entirely false. The overall truth of our view of the world thus requires not only moderate holism, but also the externalist constraint that beliefs are by nature normatively responsive to their objects, to the world that they are about. The externalist constraint, elaborated by Davidson’s metaphor of triangulation, requires that certain basic beliefs have as contents situations in the objective world, while the holistic constraint requires that the objective content and overall truth of these basic beliefs be imported into the less basic regions of our view of the world. Or, to put the point in slightly different terms, which are perhaps somewhat misleading, the holistic constraint implies a kind of coherentism, while the externalist constraint implies a kind of correspondence.
The upshot of this view would seem to be, in effect, that the distinction between appearance and reality, essential to any formulation of radical skepticism or conceptual relativism, has no place at the most basic level of belief or perception. If the world so much as appears to us to be a certain way, that can only be because, at this more basic level, appearance and reality must be one and the same. Hence we are in “unmediated touch with the familiar objects whose antics make our sentences and opinions true or false” (VICS, 198). This conclusion undermines not just skepticism but conceptual relativism. On this view, language cannot be characterized as “a screen or filter through which our knowledge of the world must pass”. So there could not be, in particular, radically different conceptual screens or filters.

IV

In the preceding sections I’ve summarized Davidson’s treatment of the scheme idea and my evaluation of his arguments. Davidson explores the idea of a conceptual scheme indirectly, by considering the idea of a difference in schemes, which he takes to involve two conditions that cannot be jointly satisfied: conceptual schemes are supposed to differ in relation to a common entity, and yet interpretation or translation by reference to that entity is supposed to be impossible.

This strategy builds on his initial observation that “the dominant metaphor of conceptual relativism” is paradoxical, in that different “points of view” can be understood by reference to a “common co-ordinate system” (VICS, 184). So, in other words, it would seem that different points of view could, in principle, be understood from a single point of view. In arguing for this conclusion, Davidson examines ways in which the scheme/content relation has been characterized—that is, the two main groups of more detailed metaphors, metaphors of organization and fit, that are supposed to render the dominant metaphor of points of view more precise and defensible. The first group of metaphors concerns reference and the second concerns truth, and neither of these concepts coheres with the claim of incommensurability.

Donald Davidson, “Introduction”, in Inquiries into Truth and Interpretation, xix.
We’ve seen that there are also more general considerations to do with the nature of interpretation, meaning and language learning that, in Davidson’s view, rule out the possibility of a difference in scheme. The latter arguments do not directly address any particular account of the scheme/content relation, but suggest that there could not be any such account, given the nature of language. Taken together, we then have a range of destructive arguments intended to show that a certain traditional philosophical conception of mind and world, subjective and objective, is untenable, and also a more constructive range of arguments that indicate the alternative picture of mind and world that Davidson defends.

I’ve argued that, with some important qualifications, these arguments constitute a powerful case for a conclusion somewhat weaker than Davidson’s: there could not be alternative schemes, in the sense that any two languages, theories or systems of concepts trained on a single world must be such that understanding is possible, at least in principle. In the limiting case, understanding would consist simply in construing both of any two putatively divergent schemes as fragments of a single, more comprehensive language or system of concepts, or as theories that are either consistent or inconsistent accounts of a single world. The possibility of integrating languages, theories or systems of concepts in this way is not the same as the possibility that actual human beings come to understand any language, theory or system of concepts. Nor is the latter possibility implied by the version of Davidson’s argument that I have defended. But, in my view, the former possibility is enough to establish that talk of conceptual schemes is empty.

In this section, I’ll discuss in more detail these differences between my position and Davidson’s, and recap the reasons for my rejection of certain aspects of his discussion. First, I should mention some important respects in which I think Davidson mischaracterizes the nature and sources of the scheme idea, at least on occasion. As we’ve seen, Davidson describes the scheme/content dualism as the result of a certain line of thought based in rejection of the analytic/synthetic distinction. The idea of empirical content is originally introduced by way of a contrast between analytic and synthetic truth, i.e., “a distinction within language between concept and content” (VICS, 187). Davidson thinks that the dualism of the analytic and the synthetic is not conducive to conceptual relativism, but it plays a key role in his account because the synthetic sentences of a
language are supposed to have empirical content. This notion of content can be retained in the absence of analyticity. Empiricists may then be led to replace the analytic/synthetic dualism with “a dualism of total scheme (or language) and uninterpreted content” (VICS, 187).

I’ve argued that this characterization of the dualism of scheme and content is too narrow. For supposedly the “empirical content” with which schemes are to be contrasted may be identified with such entities as “the facts” or “the world” (VICS, 189). A contrast between languages or theories, on the one hand, and the facts or the world, on the other, does not require empiricism. Likewise, Davidson links the scheme/content to Cartesian skepticism, a view that need not involve empiricism. And, as we’ve seen, there appear to be views that, on Davidson’s account, at least, embody the scheme/content distinction but are inconsistent, or at least in considerable tension, with empiricism.

So although empiricism is one important instance of the more general dualism of scheme and content, what is really at issue here would seem to be the idea that the subjective can be sharply distinguished from the objective—e.g., that beliefs, perceptions or meanings of utterances do not depend for their content and identity on the extra-mental or extra-linguistic world. Of course, if that were true, it would seem to be possible that one person’s states of mind or utterances could be inaccessible to an interpreter, since the interpreter’s only access to these is by way of what he can learn about the speaker’s environment and, by hypothesis, the environment is not sufficient to determine those contents. (There are other ways in which the idea that the subjective and the objective are independent could be illustrated or explained, of course.)

Not only is Davidson’s account of the sources of conceptual relativism or scheme/content dualism overly narrow. In my view, it is also confused as concerns the relation of the analytic/synthetic distinction to the scheme/content distinction. Although it is not clear from his brief discussion of the relation of these distinctions (or purported distinctions), it appears that, on Davidson’s account, the notion of analyticity is inconsistent with the scheme/content distinction. For instance, in summarizing his discussion of the possibility of a total failure of translation, Davidson writes that “a fixed stock of meanings” does not provide “a ground for comparison of conceptual schemes” (VICS, 195).
This seems to be a reference to his earlier remark that the analytic/synthetic distinction requires “a fixed system of concepts” by means of which different possible worlds can be described: “the clarity of the contrasts between worlds … depends on supposing our scheme of concepts, our descriptive resources, to remain fixed” (VICS, 187). Again, it is not clear what exactly Davidson means in making this point about talk of possible worlds, but it seems that he takes the distinction between “concept and content” associated with the analytic/synthetic distinction to somehow rule out the possibility of incommensurable schemes (or “concepts”, “meanings”, “descriptive resources”, etc.).

But why would that be? On the face of it, it would seem possible that different languages might be associated with different, incommensurable systems of meanings or concepts, each system being determined or constituted by the analytic truths of the language. The mere fact that meanings, concepts or schemes of description must be fixed, on this view, does not imply that they must be universal, i.e., that all possible languages must be associated with the same set of analytic truths. Unless we can assume that there is only one scheme of concepts, or that all schemes can be understood by reference to some single, basic scheme, the analytic/synthetic distinction would seem to be consistent with conceptual relativism. (Assuming, that is, that conceptual relativism is internally consistent.)

V

I now turn to the most important difference between the thesis that I’ve defended and Davidson’s. Although his initial discussion of conceptual schemes may be open to different interpretations, later writings, at least, make it clear that his considered position is that there could not be a language that we, i.e., actual human beings, would be unable to translate or interpret. This seems to me an indefensible position, but also one that is ultimately irrelevant to the question of whether there could be incommensurable languages, theories or systems of concepts. I’ve already explained my view of this matter, but it is worth exploring in more detail some of the issues in this vicinity.
In “Seeing Through Language”, Davidson writes that we cannot really “imagine a culture where creatures communicate in ways we are permanently disabled from penetrating”:

Speculating on this possibility hardly advances the case until we decide on our criteria for communication. Fluent exchange of information, purposeful interaction? But how are these manifested? The only ends we recognize are the same as, or analogous to, our own. Information as we know and conceive it has a propositional content geared to situations, objects and events we can describe in homespun terms.\textsuperscript{178} It seems clear, then, that in denying that there could be alternative “conceptual schemes”, Davidson means not only to deny that there could be languages \textit{incommensurable} with our own, but also to deny that we can make sense of the idea of a language or form of communication that we could not describe “in homespun terms”.

His reasoning seems to be that the concepts of language, communication or propositional content are intelligible only in relation to the particular “situations, objects and events” that figure in our communication. That is, this claim appears, here, to be a premise for his conclusion that alternative schemes are impossible, not just an alternative formulation of that conclusion. The claim that there could not be languages that we would be unable to interpret is fairly explicit in “On the Very Idea of a Conceptual Scheme”, for instance in Davidson’s denial that we can “make sense of there being a language we could not translate at all” (VICS, 192). However, it is important to note that the argument or arguments advanced there are quite different from the argument suggested in the passage cited above.

The latter argument takes as a premise the claim that the very concept of communication or language (and related concepts as that of “fluent exchange of information”, etc.) depends for its intelligibility on what we can describe or think about in “homespun terms”. But the arguments advanced in “On the Very Idea of a Conceptual Scheme” do not, on my reading, at least, make use of any premise to that effect. Instead, Davidson considers key features of language, such as reference, predication and truth, in a highly abstract manner. The arguments do not seem to require the assumption that

\textsuperscript{178} Donald Davidson, “Seeing Through Language”, in \textit{Truth, Language and History}, 128.
these concepts are given content only in relation to familiar objects or situations. For example, Davidson argues that referential differences between languages are intelligible only if we suppose that there is “an ontology common to the two languages” (VICS, 192). The argument seems to be directed (as I think it should be) at the idea that reference, considered in abstraction from the particular referents of our terms, could provide a basis for the idea of a difference in scheme.

So construed, however, the argument then allows that reference can be understood in some way that does not depend essentially on our knowledge of some *particular* ontology, i.e., the ontology of familiar objects, events and situations. Davidson’s point seems to be that although we can make sense, in the abstract, of differences in reference, such differences conceptually require a common ontology and hence the possibility of describing the differences against “a background of generally successful translation” (VICS, 192). For instance, he says merely that in such cases, different languages are associated with “concepts that individuate the same objects” (VICS, 192)—not that the concepts must be ours, or that the objects must familiar ones. And he begins his argument, recall, by pointing out that the concept of organization “applies only to pluralities” (VICS, 192). But none of these steps in the argument would be necessary if we could assume or show that the *very idea* of reference is just the idea of reference to the particular objects and kinds of objects that figure in our communication.

I think similar points can be made concerning other important subordinate arguments in Davidson’s discussion of the scheme idea. For instance, one might suppose, in line with Davidson’s later remarks about the concepts of communication, information and the like, that the concept of *truth* depends essentially for its content on the particular statements that we take to be true. Of course, this would imply that no theory could be both true and yet untranslatable into our language, since the truths stated by the theory would be identical with certain statements of ours. But Davidson does not seem to argue in this way. Instead, he makes a highly abstract argument from the relation between the concepts of truth and translation. I’ve objected to this argument, but what matters for present purposes is that the argument would surely be unnecessary if the very idea of a true theory were essentially just the idea of a theory that expresses our view of the world.
Finally, Davidson’s argument for the necessity of triangulation in thought and communication does not presuppose or imply that the only entities in relation to which speakers can triangulate their responses are, say, “knives and forks, railroads and mountains, cabbages and kingdoms” (VICS, 192). Rather, the argument appeals to very general features of communication. There must be some way of fixing the object about which a speaker is saying something, Davidson argues, and the only way in which this can be fixed is by identifying the cause of his utterance with some object, situation or event that is also the cause of responses in other speakers. This line of thought seems persuasive, but makes no reference to any particular kind of object, or even to any particular kind of speaker. So we cannot conclude that the very idea of communication, and hence of triangulation, somehow depends on our knowledge of familiar objects.

VI

Davidson’s treatment of the scheme idea appears to allow, at least for the sake of argument, that we can have some notion of language, communication, truth, etc., abstracted from the particulars of our own language or view of the world. At least, to the extent that the arguments succeed, they cannot be taken to rest on the assumption that these notions are only intelligible in relation to familiar contents, information, objects or situations. But it might be that Davidson’s claim about the dependence of these key concepts on their familiar applications is not intended to introduce a different argument for the impossibility of an uninterpretable language. Perhaps this claim is intended to be an alternative statement of that conclusion, or a further conclusion.

In any case, it is worth considering more closely what exactly Davidson means to deny as regards the idea of an alternative scheme. In denying the possibility of “a language we could not translate at all” (VICS, 192), one might deny the possibility of any (or all) of the following:

1. a language geared to a radically alien range of objects, situations or events;
2. a language that we would be unable to interpret or translate;
3. a language that we would be incapable of interpreting or translating;
4. a language incommensurable with our language.
I’ve argued only that (4) is not a real or intelligible possibility, and that Davidson’s arguments are unsound unless they are taken, where possible, to be arguments for that conclusion. In the passage cited earlier, Davidson seems to endorse the far stronger thesis that (1) is not a real or intelligible possibility, which implies that none of (2)-(4) are possible. I’ve already argued that, in my view, Davidson’s arguments cannot establish this extremely radical conclusion, because they are pitched at too high a level of abstraction. If Davidson takes the denial of the possibility of (1) to be an alternative statement of his conclusion that there could not be a language that we are unable to interpret, that is a mistake. The conclusion is equivalent to the denial of the possibility of (2), not (1).

Finally, the denial of the possibility of (1) cannot be legitimately used as a premise in an argument for Davidson’s conclusion, as it appears to be in the passage from “Seeing Through Language”, because it is a much stronger and more controversial claim than his conclusion. If the denial of the possibility of (2) is not self-evident, and so should “emerge as the conclusion of an argument” (VICS, 186), the denial of (1) should emerge as the conclusion of that same argument, whatever it might be, together with some additional argument. So if Davidson means to deny the possibility of (1), he has not given an argument for that radical position, and cannot reasonably appeal to that position as the basis for concluding that (2) is not possible.

In any case, I would argue that (1) is a clearly intelligible possibility. For example, we know that many of the objects, situations and events that we think and communicate about would seem very unfamiliar to our distant ancestors—chromosomes, computer programs, stock markets, and so on. To be sure, there are also many other kinds of objects, situations and events common to our thoughts or communication and theirs. Prehistoric hunter-gatherers presumably talked about the weather, about animals, rocks and trees, etc. But it does not seem to be necessary that there is this common range of content, or common range of objects, situations and events. Intuitively, it would seem that there could be some alien species whose communication concerned only objects, situations and events totally unrelated to the subject matter of ours. After all, it is surely a biological accident that the sense organs and nervous systems of language users are, so far as we know, sensitive to certain kinds of objects, qualities or events rather than others.
This intuition can be defended by reflection on the kind of conceptual or linguistic difference under discussion, a totally innocuous and familiar kind of difference. As I’ve argued in earlier chapters, the kind of difference in content, information or subject matter that I claim to be possible is simply a difference between “languages that have evolved in distant times and places” and hence “differ extensively in their resources for dealing with one or another range of phenomena” (VICS, 183-4). The Inuit may talk about snow and ice, while people in the tropics talk instead about coconuts and palm trees. If neither group has ever been to the other’s environment, or heard about it, their languages may differ extensively. This much is obviously and trivially true.

But this is precisely the kind of difference that I am saying is possible between speakers or thinkers whose perceptions or external environment are radically unlike ours. Suppose that some alien species navigates by sensory perception of electromagnetic fields, or sonar, or some still more exotic form of perception, but have no sense of sight, hearing or taste. If such creatures could triangulate on inter-subjectively accessible objects or aspects of their environment—and surely it is logically possible that they do so—they could communicate about that environment. But presumably the contents of their utterances would be very alien to us.

Alternatively, it would seem to be logically possible that some species perceives only objects much larger or much smaller than those that we normally talk about. If they speak only about sub-atomic entities, for example, the contents of their communications would not be describable in familiar terms. Or we can imagine creatures whose lifespan is vastly longer than ours, with the result that the situations and events they talk about are beyond our normal perception or understanding—events in the evolution of a species, say, or the formation of galaxies.

These scenarios may be fanciful, but they seem to be intelligible, and they do not involve any kind of conceptual or linguistic difference over and above the innocuous kind of differences that we can observe between actual languages appropriate to different environments, different subjects or interests. I don’t think there is any reason to deny that these actual differences could not be significantly greater than they are, as in the examples given just now.
Notice also that these possibilities are not equivalent to (2), i.e., the possibility of a language that we would be unable to interpret. For it might well be that we could interpret any of these logically possible languages, just as, given the right circumstances and perhaps some ingenuity or imagination, the Inuit could interpret languages that evolved in Africa or India. Maybe we would be able to interpret a language geared to sub-atomic objects and events if we had a correct theory of sub-atomic physics and some sophisticated technology.

This highlights again the radical nature of the claim that (1) is not an intelligible possibility. That claim is far stronger than Davidson’s conclusion in “On the Very Idea of a Conceptual Scheme”, that we cannot make sense of “there being a language we could not translate at all” (VICS, 192). To deny the possibility of (1), as Davidson appears to do in “Seeing Through Language”, is to deny even that there could be a language that, while *translatable* into ours, was simply *very different* from ours, i.e., a language geared to objects, situations and events other than those that we actually think and speak about.

What about the weaker thesis that (2) is not a real or intelligible possibility, i.e., that there could not be a language that we would be unable to interpret? In Chapter 3, I distinguished between ability and capacity: we are able to interpret a language if it is possible that we interpret it, whereas we have the capacity to interpret a language if it is possible that we able to interpret it. So, to repeat the example given earlier, a person may be unable to interpret a Chinese utterance, but capable of interpreting it because he could study Chinese and then be able to interpret that utterance. By contrast, he might lack even the capacity to interpret certain statements about, say, quantum mechanics. No matter how long he tries to understand quantum mechanics, he will never understand it; the concepts involved transcend his intellectual capacity.

Now it seems clear enough that there could be languages we are unable to interpret. We can imagine that Martians speak a language associated with scientific theories vastly more advanced than any that we know. In order to interpret the language, let us suppose, we would need to undergo decades of study. This is surely something we can imagine. And although, as I’ve emphasized, it is hard to say what exactly Davidson’s
position is regarding this kind of scenario, it seems likely that he would allow it to be an intelligible possibility.

At any rate, I think that this clearly is intelligible, for reasons already given. We need only imagine a difference between our language and some other, evolved in a very different environment, or on the basis of very different kinds of perceptions, i.e., a difference of type (1). But then, to make sense of the idea that we are unable to interpret that language, we need only imagine that we stand to it in something like the relation of the relation of the Inuit to an African or Indian language (for example). We are not able to interpret it, but given the right circumstances, enough time, patience and ingenuity, we could acquire that ability. The scenario depends only on innocuous assumptions about language and interpretation, which of course depend on the environments within which they operate.

So both (1) and (2) seem to be intelligible possibilities, and neither has any bearing on the idea of a difference in scheme. But, as I’ve said, it is not entirely clear what Davidson’s position is regarding these possibilities. It is in relation to (3) that my position clearly differs from Davidson’s. I’ve argued that there could be languages that we are **incapable** of understanding or interpreting: our cognitive relation to these languages is like the relation of the least intelligent actual people to the specialized languages of, say, mathematicians or physicists. Such people will never acquire the ability to understand or interpret those languages (or language fragments), regardless of time and effort.

Since we know that this relation actually obtains between certain speakers and languages, it seems natural enough to suppose that it could obtain between all human speakers, even the most intelligent or perceptive, and some other language. We might lack the capacity to interpret certain languages because the changes we would have to undergo in order to be able to interpret them would result in creatures that could not reasonably be counted as human beings. By analogy, squirrels lack the capacity to understand Chinese because, even though squirrels could perhaps evolve into creatures able to do so, those creatures would not count as squirrels by any reasonable standards. This is the kind of relation that I think is logically possible, at least, between actual human beings and certain languages or groups of speakers.
This is not the only way in which this kind of relation might obtain. Again, radical differences in sensory modalities, or in psychological or behavioral dispositions, could also result in our incapacity to interpret certain languages or forms of behavior. Davidson denies that this is a possibility. The thesis that we can’t “make sense of there being a language we could not translate at all” (VICS, 182) must mean, at least, that we can’t make sense of there being a language we are incapable of interpreting. My reasons for rejecting Davidson’s position are then that, first, it seems to me that I’ve given a fairly clear account of how this kind of scenario might come about, which suggests that the idea is intelligible, and, second, the arguments regarding the intelligibility of the scheme idea that I find convincing do not support Davidson’s thesis. So there is a prima facie reason to believe that the idea of “a language we could not translate at all” is intelligible, and I’ve found no good argument against this belief.

VII

In this connection, it is worth asking whether Davidson means to deny that we can make sense of the idea that, concerning some particular “form of activity”, that form of activity is at once not interpretable and yet a kind of “speech behavior” (VICS, 185) or whether, instead, he means to deny the proposition that there could be a language that we are incapable of interpreting. That is, Davidson might be taken to mean that (A) For any form of activity $x$, it is not intelligible to hold that $x$ is a language and yet not something that we are capable of interpreting, or rather that
(B) It is false that there could be a language that we are not capable of interpreting.
One might hold (A) to be true without taking a stand on the truth of (B). Perhaps the proposition that there are such languages is intelligible, and even true, but there is no situation in which we could intelligibly conceive of some particular form of activity as both speech behavior and something we are incapable of interpreting. I’ve assumed that Davidson’s argument is intended to establish (B). The reason is that, unless one argues for (B), one must allow the possibility that there are alternatives schemes; at most, one could hold that it is not intelligible to claim to have identified such
a thing. The conceptual relativist need not claim that we could intelligibly identify some particular form of behavior as language and, at the same time, as something we are incapable of interpreting. Indeed, a radical conceptual relativist might take it to be essential to his view that we can never intelligibly identify such a thing—this might be a consequence of our total imprisonment within our own, parochial scheme.

But if Davidson’s position is consistent with this kind of view, as it would be if he were to argue merely for (A), it cannot be a position that has the implications that he claims for our understanding of how language and mind are related to experience or reality. On the kind of view we are considering, it might be perfectly intelligible to say that language mediates between mind and world, shapes or filters our experience, etc. But given the wider context within which he situates his attack on the scheme idea, Davidson clearly does mean to deny the intelligibility of such a conception of language.

My contention has been that, contrary to what Davidson seems to believe, there is no need to assert (B), or some conclusion to that effect, for the purpose of denying the intelligibility of conceptual relativism or of the scheme/content distinction. While the truth of (B) would imply these conclusions, (B) is itself implausible, for reasons already discussed, and allowing the possibility of languages we are incapable of understanding or interpreting is not the same as allowing the possibility of incommensurable languages.

So, to summarize, I take the purported possibility of incommensurable languages (i.e., (4), in my earlier list) to be importantly different and far more controversial than any of the other three possibilities canvassed in this section: the possibility of languages geared to radically different objects, situations or events, of languages we are unable to understand, or of languages we are incapable of understanding. These are all logical possibilities, at least, and have none of the unwelcome consequences that Davidson associates with the dualism of scheme and content. Those are consequences of the very different and controversial claim that languages might be incommensurably different, which is to say, that there is no common measure of certain languages: no single language, theory or system of concepts by means of which the differences could be explained and understood. If that is not an intelligible possibility, the dualism of scheme and content is untenable, regardless of any facts about our abilities or capacities.
I’ve argued that there is a distinction between contingent features of our language, which happens to be geared to certain kinds of objects, events and situations, and learned on the basis of certain dispositions and powers, and other, necessary features of language that ensure the possibility of interpretation across languages. For the reasons we’ve considered in the preceding sections, it is not defensible to hold that the former features of our language are constitutive of the concept of language (or of the concept of content, information, communication, and so on). But neither do those features play any essential role in those of Davidson’s arguments that I take to be successful and relevant to the question of whether there could be alternative conceptual schemes.

As I’ve said, it is not entirely clear whether Davidson understands his argument or his general position in the way that I’ve argued against. For instance, he writes that his account depends on an assumption about “the shared propensities of people to react in observably similar ways to their environment”, and so on what he takes to be “an empirical fact about people”.\textsuperscript{179} Does this mean that it depends on the empirical fact that people have the \textit{particular} shared propensities that they do, or merely on the empirical fact that they have shared propensities? Only the latter fact, in my view, is relevant to claims about the nature of language and thought, and hence the possibility of a difference in scheme. In any case, it is worth considering this important difference between my position and a position, which may be Davidson’s, as a way of clarifying what I take Davidson’s arguments, where successful, to establish.

\textbf{VIII}

I now turn to some more general philosophical topics related to the dualism of scheme and content. First, recall that Davidson takes his repudiation of the scheme idea to imply a radical revision in the way that we conceive of the distinction between the subjective and the objective. Is this true? In my view, the most defensible version of Davidson’s position has consequences less dramatic than those that he describes. To reject the scheme idea is not necessarily to “be rid of representations”,\textsuperscript{180} for instance.

\textsuperscript{179} Donald Davidson, “Reply to Stroud”, in \textit{The Philosophy of Donald Davidson}, 165.
\textsuperscript{180} Donald Davidson, “The Myth of the Subjective”, in \textit{Subjective, Intersubjective, Objective}, 46.
Rather, Davidson’s critique of the scheme/content dualism constrains theories of representation.

The notion of mental or conceptual representation is vague and broad, but let us consider one version of the idea, according to which we perceive objects or situations in the external world only indirectly: what we directly perceive is a mental entity, an image, perhaps, or an image-like entity, which may or may not represent the environment. This kind of view may be problematic in a number of ways, but Davidson’s case against the scheme idea does not imply that it is false or unintelligible. What is implied is rather that, if there are representations in this sense, their contents cannot diverge radically from the general nature and structure of what they represent, and hence cannot diverge radically from the general nature and structure of other representations.

The reason for limiting the consequences of Davidson’s position in this way is that representationalist theories of the kind under discussion are, in some sense, or in part, theories about the mechanisms of thought and perception. But Davidson’s argument against the dualism of scheme and content does not tell us anything about these mechanisms. At most, it shows that they cannot be understood in ways that would suggest or imply the possibility of incommensurable ways of representing the world. Even if that is so, it could still be a psychological fact, seemingly, that our experience of the world is mediated by mental images, or something of that kind.

More precisely, what is ruled by Davidson’s argument against the scheme/content dualism is a certain kind of mediation. Davidson says that he rejects any view that places “intermediaries between our beliefs and their objects in the world”, but he is careful to add that it is really *epistemic* or *semantic* intermediaries that he takes to be the source of trouble. For there are obviously causal intermediaries between our beliefs and their objects, and his account of communication and thought can accommodate these. For instance, sensory stimulation is obviously a causal intermediary between the belief that one is seeing a rabbit and the rabbit itself. Davidson distinguishes his position from Quine’s by allowing that the belief is caused, in some complicated way, by sensory

---

stimulation (among other things), while denying that stimulation is a kind of evidence for the belief, or the basis somehow of the content of the belief.

I think the same distinction between causal intermediaries, on the one hand, and epistemic or semantic intermediaries, on the other, can be made as concerns the notion of representation. To be sure, it is much less obvious that representations, in the present sense, play any kind of causal role in our knowledge of the world. But even if they do, Davidson’s attack on the scheme idea would seem to accommodate this causal role; what is ruled out is merely the kind of theory that would assign to these representations an epistemic or semantic role, rather than a merely causal role. I would argue that Davidson’s position regarding scheme and content is even consistent with theories that posit sense data, so long as these are not taken to play an epistemic or semantic role.

In fact, I would suggest that a rejection of the scheme/content distinction is consistent even with the stronger claim that these entities, mental images or sense data, play epistemic or semantic role in our knowledge of the world in a way that does not conflict with Davidson’s rejection of the scheme idea. For it would seem to be logically possible, at least, that our knowledge or experience and the meaning of our sentences is mediated by these kinds of entities, but that, because of Davidson’s arguments, any such form of mediation must be commensurable with any other. That is, suppose that when I believe that there is a rabbit in the garden, I am directly related only to some complex entity having sense data as constituents. The content of my sentence “There’s a rabbit in the garden” is somehow determined by that entity, and that entity is also something that I take to be evidence for further beliefs.

This kind of view would be inconsistent with Davidson’s position only if the mediating entities “could be just as they are and the world be very different”. But we could stipulate that they could not be as they are, or what they are, unless the world that they represent is basically and generally the way that it is represented. We could hold that the character of these representations is determined by the external world, and that beliefs, by extension, are also determined in that way, although only indirectly. With this stipulation, it is not possible that we could be in massive error about the world, or that there could be incommensurably different ways of representing the world.

---

182 “Myth”, 43.
Perhaps this point deserves clarification. If beliefs or perceptions are mediated by representations, it is logically possible to have the representations in the absence of what they represent. One could have mental images of a rabbit in the absence of any rabbit anywhere in the world. In this sense, representations “could be just as they are and the world be very different”. This follows merely from the assumption that representations are distinct from what they represent. But could these representations play the epistemic and semantic role that, on the present view, they play in our mental lives, even if the world were utterly unlike what we take it to be? Alternatively, could there be incommensurable representations, that play a similar role in some alien form of thought, but which are nevertheless representations of the actual world?

Seemingly, one might hold that our representations actually represent the world only because they are linked, by triangulation—a more complex form of triangulation than Davidson describes—to inter-subjectively available objects and situations. Presumably those same representing or mediating entities, if there are such things, could exist in a very different world, perhaps even a world in which there are no believers or perceivers. But if Davidson is right about scheme and content, they could not, in a radically different world, play the epistemic or semantic role that they actually play. (Some other, radically different range of representations would be needed, in that world.)

Again, I am not trying to argue for any such view, and in fact I am highly skeptical of the kinds of views under discussion, for various reasons. My point is merely that one could perhaps devise a theory of representation, or something like it, that is consistent with Davidson’s treatment of the scheme/content dualism, which assigns to representations a causal and an epistemic or semantic role.

For similar reasons, it would seem to be consistent with Davidson’s position, on my reading, to allow that we represent the world by way of propositions (however exactly these are to be construed). Again, what is not intelligible, if Davidson is right, is the idea that there could be incommensurably different propositional representations, i.e., that the propositions that mediate between one person’s environment and another differ in some special way that would prevent understanding or interpretation. But it might well be a psychological or metaphysical fact that in believing that there is a rabbit in the garden, one is related to the rabbit (or his being in the garden) only indirectly, by way of
a certain abstract object—a set of possible worlds, perhaps, or a Fregean Thought. Of course, if we represent the world by some relation to propositions, these intermediaries would be both semantic and epistemic. But perhaps one might hold that, for reasons we’ve considered, all propositions must be commensurable, and that we have access only to those propositions that can be grasped (somehow) on the basis of triangulation.

So, to summarize, it seems to me that the very idea of representation or mediation, of objects before the mind, does not clearly or directly imply a dualism of scheme and content, as Davidson seems to think. Some such notion is a necessary condition for the dualism, and for conceptual relativism, but it is not sufficient. Even if the entities said to represent the world or mediate between beliefs and their objects are assigned an epistemic or semantic role, this is still not sufficient for conceptual relativism. The case against the scheme/content distinction constrains the notion of representation, rather than revealing it to be empty or unintelligible. Some further argument would be needed for that conclusion.

We have already considered one such argument, which Davidson takes it to be part of his larger case against conceptual schemes: the slingshot argument. Davidson takes the notions of representation or mediation and correspondence to fact to be very closely related, perhaps ultimately identical. The connection between representation and correspondence is clear enough if the representations are propositions, or entities with propositional content. For what could those entities represent, if not facts, or some other kind of truth maker? If the slingshot argument shows that there is no way to individuate facts, that notion of representation, at least, would be bankrupt. As I’ve argued, it is less clear that the notion of correspondence to fact has any connection to the idea of a difference in scheme. So the slingshot argument, if sound, shows that a certain kind of representationalism is untenable, for reasons independent of Davidson’s attack on the scheme idea.

Now it may be that there are good independent arguments that rule out every form of representationalism. In that case, we ought to accept Davidson’s surprising contention that “beliefs may be true or false, but they represent nothing”183. However, I’ve suggested that there are reasons to doubt the intelligibility of this contention. Truth and

183 “Myth”, 46.
falsity, and hence belief, seem to me to be essentially representational phenomena. In some sense, which I have not tried to articulate in a rigorous or precise way, to have a belief just is to represent the world as being a certain way; the belief is true just in case the world is as the belief represents it to be. And if this is so, then regardless of how exactly we understand this essential feature of belief and truth, there is some contrast to be drawn between representations, or acts of representation, perhaps, and what is represented. To reject any such contrast, any notion of representation, would require us to give up the concepts of belief and truth.

My main reason for objecting to this aspect of Davidson’s philosophy is simply that on the very modest assumption that knowledge of truth conditions is necessary for linguistic understanding and meaning, this knowledge can only be seen as a kind of representation, in some sense of the term. To know that “Snow is green” is true if and only if snow is green is to have in mind a way that things could be in the world—to have it “before the mind”, we might as well say. In this case, what one has in mind is not how things actually are.

But then, however exactly the distinction between what one has in mind and how things are should be understood, there is some distinction here. On the one hand, there is the condition in mind, or one’s state of mind in knowing the condition, and on the other hand, there is the world: snowflakes, white and green things, and so on. Unless there is some sense in which one represents the world by means of the truth condition, or represents it in knowing or thinking about the condition, it will be unintelligible to suppose that truth depends on the world. Or so it seems to me. But a concept of truth that is not in some sense dependent on the world is not a concept of truth at all.

IX

The preceding section suggests that Davidson may also be wrong to infer from the unintelligibility of the scheme/content dualism to the unintelligibility of “the concept of an uninterpreted reality, something outside all schemes and science” (VICS, 198). This conclusion might be based on an overly simplistic view of the relations between the concepts of correspondence, representation and conceptual scheme: to reject any of these
is to reject the others, and with them the concept of an “uninterpreted reality … outside all schemes and science”, since that concept is intelligible only by way of some contrast with schemes.

In my view, one might well maintain some notion of representation while rejecting talk of correspondence or conceptual schemes. Indeed, as I’ve said, I think that one is forced to accept some such notion on pain of rejecting the very concept of truth or belief, and also that of meaning, if meaning involves truth conditions. But if it makes sense to say that sentences or beliefs represent the world in some sense, i.e., that they represent objects or arrangements of objects, the concept of a reality “outside all schemes and sciences” can survive Davidson’s arguments against the scheme/content dualism.

It is unclear what Davidson means to say in the passage cited just now, but I will hazard some further speculations. Davidson’s reasoning seems to be roughly as follows. The idea of a conceptual scheme is, essentially, the idea of a subjective point of view or way of representing the objective world, an “uninterpreted reality”. In rejecting the dualism of scheme and content, we thus reject any such contrast between subjective and objective. Specifically, we must give up the idea of a reality that lies “outside all schemes and science”.

Thus, Davidson writes:

It would be wrong to summarize by showing how communication is possible between people who have different schemes, a way that works without need of what there cannot be, namely a neutral ground, or a common co-ordinate system.

(VICS, 198)

So, on this account of his treatment of the scheme idea, it would seem that the conclusion of the argument is that we cannot make sense of the idea of an alternative scheme because there can be no “common co-ordinate system”—that is, an “uninterpreted reality”—in relation to which schemes could differ. But does it make sense to say that reality, or what we take to be reality, at least, is in some sense an *interpreted* reality? This would seem to be essentially a way of saying that reality, or what we take to be reality, is itself a product our conceptual scheme, and that we cannot imagine others because we are so thoroughly confined to our own.
Davidson’s argument, on this reading, would seem to be directed less against the idea of a conceptual scheme than, instead, the idea of a mind- or language-independent reality. We might put the point this way. The conceptual relativist puts forward an unstable thesis that combines elements of realism and idealism. On the one hand, there is a “common co-ordinate system” in relation to which schemes differ. This would seem to involve a kind of realism, a commitment to some mind- or language-independent world. On the other hand, there is no way to interpret across schemes, because, supposedly, one’s conceptual scheme shapes or filters one’s experience and conception of the world in some profound and pervasive manner. This element in conceptual relativism seems to involve something like idealism, since the idea is that reality is somehow determined by the mind or by language.

Davidson then tries to dismantle this hybrid position by arguing against the first, realist element, and arguing for a more consistent and complete form of idealism. If we are consistent idealists, there is no logical space for relativism, because there can be nothing outside our “scheme” that others might represent in different ways. Moreover, a consistent idealist will be unable to make sense of idealism, ultimately, because the thesis that reality is ideal, or mind-independent in some sense, is only intelligible if we think that reality could be independent of the mind—i.e., of our language and the general picture of the world associated with it. Once we see that that latter idea is simply unintelligible, the claim that reality is ideal becomes unintelligible as well.

Perhaps this is why Davidson argues from his claim that there can be no “neutral ground” or “uninterpreted reality”, which seems to express a kind of idealism, to the seemingly realist conclusion that without the dualism of scheme and content, we “re-establish unmediated touch with the familiar objects whose antics make our sentences and opinions true or false” (VICS, 198). We are left simply with objects, which are neither “real” nor “ideal” in any traditional sense, because the contrast is empty.

This may be how Davidson understands the course of his argument against the scheme/content distinction. However, this account of the argument does not seem to be consistent with the details of the subordinate arguments that he deploys against the metaphors of conceptual relativism. For instance, in arguing against the idea that schemes are systems of organizing principles, he emphasizes that differences in reference
require “a common ontology” (VICS, 192). But a common ontology, which is not to say a common set of concepts, would be precisely a “common co-ordinate system” or “neutral ground” by means of which putative differences in scheme could be understood.

The argument, recall, was that one could make sense of differences in extension by comparing both languages with a single plurality of objects. Surely the objects would be “something outside all schemes and science” (VICS, 198), or outside those under discussion at least. If languages organize cabbages and kingdoms in different ways, the cabbages and kingdoms can be distinguished from the concept of a cabbage or a kingdom, or from the words “cabbage” and “kingdom”.

Similarly, in arguing against the second group of metaphors associated with scheme idea, Davidson stresses the connection between the concept of truth and Tarski’s Convention T. But if this connection is as he describes it, putatively different schemes could, again, be understood by appeal to a “common co-ordinate system”, or “something that lies outside all schemes”. In this case, the common system consists not of a shared ontology, but rather the meta-language(s) into which putatively divergent schemes can be translated. I have objected to this phase of Davidson’s argument, but the point is that the argument appears to depend on just the kind of common co-ordinate system that Davidson, in this later remark, says is not possible. My own reading of the argument, by contrast, has emphasized the ways in which Davidson’s arguments against the scheme/content dualism, where successful, reinforce his initial observation that the metaphor of points of view is intelligible, “but only if there is a common co-ordinate system on which to plot” (VICS, 184) the differences.

In any case, it is hard to make sense of Davidson’s rejection of “the concept of an uninterpreted reality” or of “something outside all schemes and science”. Surely the earth or the Eiffel Tower are “outside all schemes and science” in some straightforward and innocuous sense. It would be very counter-intuitive to hold that such things are mental or linguistic entities, or that they exist only relative to a conceptual scheme or theory. It is unclear what that could even mean. To be sure, the earth (say) is an “interpreted” entity in some sense, and something that is, so to speak, “inside” our scheme or theory or world view. People have beliefs about it, and conceptualize it in various ways, and so on.
This, at least, is the only meaning I can assign to the claim that reality is “interpreted”. But it would be invalid to conclude from the fact that our beliefs are about reality, which is thus “interpreted”, that the objects of belief or conceptualization must be constituents somehow of our world view, our beliefs or concepts. Moreover, even if we can understand that strange conclusion, which seems to me quite doubtful, it would follow that what we are thinking or speaking about is simply not, for example, the earth. The earth, if there is such a thing, is not something “inside” a language, scheme or theory.

The earth is “part of the solar system, a system partly identified by the fact that it is a gaggle of large, cool, solid bodies circling around a very large, hot star”.

Whatever exactly it means to be “inside” schemes or science, or at least not “outside” of such things, the earth is somewhere else: it is circling the sun, and under our feet. So the conclusion of this reasoning would seem to be that there simply is no such thing as the earth, or, perhaps, that the concept of the earth is not intelligible.

In what sense, exactly, does Davidson mean to deny the existence of a reality outside schemes and science? Perhaps the point is that, because of content externalism, the world itself enters into our picture of the world, our beliefs and concepts. World and world view cannot come apart because the latter is constituted by the former. If this is roughly Davidson’s reasoning, it is strange, first of all, that he states his conclusion by emphasizing the emptiness of the concept of a world beyond “schemes and science”. The real upshot of this reasoning would seem to be, instead, that the concept of a scheme or science, or a view of the world, is empty: the world view collapses, somehow, into the world that gives content to the view. On the other hand, if Davidson means that each side of the purported distinction between subjective and objective depends for its intelligibility on the other, the conclusion should be that neither the concept of a world nor that of a world view is intelligible, because the subjective side of the distinction has no content.

But I don’t see that the argument for content externalism has these radical consequences. The argument rests on the fact that, in learning a language, the learner must identify the content of the sentences he is learning with actual states of world. But even if this is right, it can hardly be that the world enters into our view of the world in

---

184 Donald Davidson, “Thought and Talk”, in Inquiries into Truth and Interpretation, 168.
such a way as to demolish any distinction between subjective points of view, beliefs or theories, on the one hand, and the objective reality at which they are directed, on the other. At least, this radical conclusion cannot be reached in a way that is consistent with Davidson’s own account of meaning, for reasons suggested earlier. Some notion of representation seems essential to the concepts of belief, truth and meaning. This notion of representation, however it is to be understood, allows us to distinguish between “schemes and science” and how things are in the world, which must be “something outside all schemes and science”, for the most part, if the concept of truth is intelligible.

But we’ve also seen that Davidson’s arguments against the dualism of scheme and content do not have any clear implications concerning the mechanisms of thought and perception. So it might be, for example, that the way in which the child comes to learn the meanings of sentences by triangulation is mediated by some kind of representation. Maybe the child correlates his own mental images with those of the teacher, by way of their shared innate standards of similarity, applied to a mind-independent object that causes the mental images of both perceivers.

Again, I am not suggesting that this is a plausible view. What matters is that it merely complicates triangulation, rather than making it impossible. In short, the argument for content externalism, if sound, really shows merely that the world must enter into our world view in the sense that it must determine the contents of some basic and essential range of beliefs and perceptions. It does not establish that the world itself constitutes those contents, i.e., that the contents, or some constituents of the contents, are simply objects in the world. That is just one way in which the contents might be determined by the world. But some more complex process might achieve the same ends. Or so it would seem, at least. Once again, some further argument would be needed to show that content externalism implies that the world constitutes these basic contents, in the present (literal) sense.

Finally, even if we were to suppose that this further claim were true, this would still not be sufficient to establish the far more radical and, to my mind, unintelligible conclusion that there is nothing “outside all schemes and science” (VICS, 198). The argument for content externalism applies only to some range of perceptually and semantically basic cases. Only in such cases is it necessary, seemingly, that the world
determines the contents of beliefs or sentences. Once we have learned a language, however, these considerations do not apply, or do not apply in the same way. A speaker whose beliefs and concepts are anchored in the objective world by an initial phase of triangulation can then, on that basis, have beliefs with contents not determined in this way. For the reason why, in the basic cases, contents must be determined externally was precisely that the pre-linguistic learner does not yet have a language or system of beliefs.

To reach the conclusion that the world determines all contents by somehow constituting those contents, one would need the additional and absurd premise that all beliefs or (meaningful) sentences are true. Otherwise, there could be no explanation of the content of certain beliefs or sentences, e.g., the belief that snow is green, or the sentence “Snow is green”. Since this is clearly absurd, some other account would then have to be given of the contents of countless non-basic beliefs and sentences, and we could at least distinguish between those features of our world view and the world they represent.

X

It may seem that, in the preceding section, I have focused on some concluding remarks of Davidson’s that should not be taken too seriously. Certainly it seems that in speaking of “uninterpreted reality” (VICS, 198) and the like, Davidson is speaking loosely and perhaps making use merely of terms associated with the scheme idea, which he does not himself take to have any concrete meaning. However, it seems to me that there is a real tension in Davidson’s position between a kind of realism and a kind of idealism, and that, however these remarks are intended, they reflect that tension. For, as we’ve seen, there are key passages in Davidson’s writings that seem to express an extremely radical anti-representationalist position that amounts to something like idealism: Davidson rejects any notion of representation, and, at times, at least, seems to hold that the concept of truth is based in its relations to the concept of interpretation or translation, and not in any relation to the concept of a mind- or language-independent reality that underpins the truth and falsity of sentences.
At the same time, there is also a clearly realist thread in Davidson’s philosophy, which leads him to reject the still more extreme position defended by Rorty, according to which the concept of truth has no intelligible relation to the concept of a mind- or language-independent reality. Davidson appears to hold that truth is dependent on how things are in the world, but also that how things are in the world, insofar as such talk can be made intelligible, is itself dependent in some non-trivial way on “human concerns, like language, belief and intentional action”.\(^{185}\) (Of course, there are trivial ways in which truth depends on such things, or happens to depend on them. If there were no speakers or believers, there would be no vehicles of truth; and, as it happens, the only speakers are human, so far as we know.)

In arguing against the dualism of scheme and content, Davidson does not actually defend a kind of idealism. Rather, on one reading, at least, he argues for a kind of conceptual relativism so extreme that talk of conceptual schemes is rendered unintelligible. We have no conception of any “neutral ground” (VICS, 198) with which schemes could be compared, so the idea of a difference in scheme is empty. In effect, we are so thoroughly confined to our own point of view that we cannot even imagine what it would be to have some other. Idealism (i.e., the repudiation of any scheme-independent world) is thus a dialectical position in the argument that is ultimately abandoned, for lack of an intelligible contrast, in favor of a final position that purports to transcend all traditional accounts of the relation of mind to world.

Ultimately, then, it is not clear whether the form of realism that I’ve defended differs in any substantive way from Davidson’s view that our beliefs and sentences are made true by the behavior of “familiar objects” (VICS, 198). But this view is not consistent with the quasi-idealistic or anti-realist position that Davidson sometimes seems to express—unless we take seriously Davidson’s denial that there is something “outside all schemes and science”, which would then conflict with the reassuring claim that we are in touch with familiar objects. There is an instability here that, in my view, should be resolved in favor of a modest and natural realism that survives the rejection of scheme and content because it is built into the concepts of truth, belief and meaning. Moreover, such realism provides a better and clearer basis for the case against the dualism.

\(^{185}\) Donald Davidson, “Truth Rehabilitated”, in \textit{Rorty and His Critics}, 73.
So, to summarize my discussion in this chapter, I hold that, with some important changes and qualifications, Davidson’s strategy against the dualism of scheme and content succeeds, although it does not establish conclusions as radical as Davidson’s. Nevertheless, the argument provides a powerful and, in many respects, a highly original and thought provoking case against conceptual relativism and radical skepticism and, by extension, against traditional views of the relation of mind and world that underwrite these positions. Of course, there is nothing particularly new in the observation that the relativistic metaphor of a “dramatic incomparability” (VICS, 186) between points of view within a single space or world seems paradoxical. But Davidson’s treatment of the scheme idea amounts to a detailed examination of the ways in which central features of conceptual thought and language make it impossible to elaborate this metaphor in a defensible way.


Byrne, Alex. “Soames on Quine and Davidson”: http://web.mit.edu/abyrne/www/soamesonQandD.pdf


