

Introduction: Davidson's philosophical project

Ernie Lepore and Kirk Ludwig

Donald Davidson's work has had a pervasive influence on analytic philosophy throughout the last half century. His contributions lie primarily in the theory of meaning, the philosophy of mind and action, epistemology, and metaphysics. This volume focuses on themes connected with his work in the theory of meaning, philosophy of mind, and epistemology. The introduction first provides an overview of Davidson's work. We pay special attention to his work on truth-theoretic semantics and its integration with the project of radical interpretation, his conception of the nature of mental states, and his account of our knowledge of things around us and of our own minds. This will set the stage for a brief overview of the more detailed examinations of particular themes connected with Davidson's work that follow.

1 Truth-theoretic semantics

In a series of papers in the 1960s and 1970s, beginning in 1965 with "Theories of Meaning and Learnable Languages" (Davidson 2001a) and in 1967 with "Truth and Meaning" (Davidson 2001c), Davidson introduced and defended one of the few really novel approaches to the theory of meaning in the latter half of the twentieth century. Central to his proposal was the suggestion that an axiomatic truth theory modeled after the truth definitions that Tarski showed how to construct for formal languages could be used to give what he called "a constructive account of the meaning of the sentences" in a natural language (Davidson 2001a, p. 3). The exact import of Davidson's suggestion and whether he took himself to be pursuing a traditional project in a novel way, a reduction of meaning to truth conditions, or urging a reform of the aims of semantics, has been a matter of controversy. For example, the paper by Gary Ebbs in this volume takes issue with the interpretation offered in our 2005 book (Lepore and Ludwig 2005) over whether Davidson was pursuing a novel approach to illuminating what it is for words to mean what they do or engaging in a project of Carnapian explication (see also the exchange in (Soames 2008; Lepore and Ludwig 2011)). For

purposes of initial orientation, we will sketch the account of Davidson's project we offered in (Lepore and Ludwig 2005) and return later to the interpretive issues in discussion of contributions to the volume.

Davidson's starting point in the theory of meaning is the observation that natural languages must be compositional. By this we mean that they divide into semantically primitive expressions and semantically complex expressions that are understood on the basis of the semantical primitives and their mode of combination. Davidson argued that it was a condition on natural languages being learnable by finite speakers that they be compositional and have a finite number of semantical primitives (Davidson 2001a, pp. 8–9). This is perhaps a natural observation, but it is one which Davidson pointed out was often violated by extant proposals for analyzing various natural language locutions (*ibid.*, pp. 9–13). The observation that natural languages are compositional gives rise to the requirement that any theory of a natural language should give an account of the roles of significant expressions in the larger expressions in which they are used. The question is how to present such an account.

Prima facie, such an account should start with axioms for semantically primitive expressions a language L that (i) give their meanings and (ii) enable one to derive for each sentence s in L a theorem of the form (M).

(M) s means that p

A natural approach, looking back to Frege (Frege 1997), is to introduce entities to serve as the meanings of expressions with which they are associated. The meanings associated with complex expressions can then be treated as constructed out of the entities associated with primitive expressions, and expressions of the form “that p ” would be taken to refer to the meanings of sentences. Davidson took a dim view of this, for reasons it would take us too far afield to detail (Davidson 2001c, pp. 17–22; Lepore and Ludwig 2005, ch. 3), and he suggested that formulating a logic to deal with substitutions into the position of “ p ” in “ s means that p ,” when it is treated intensionally, would require already solving the problem of giving a compositional meaning theory, since such substitutions would have to be licensed on the basis of judgments about sameness of meaning over a class that includes complex expressions.

At this point, Davidson suggested a radical break with traditional approaches. As Davidson put it:

Anxiety that we are enmeshed in the intensional springs from using the words “means that” as filling between description of sentence and sentence, but it may be that the success of our venture depends not on the filling, but on what it fills. The theory will have done its work if it provides, for every sentence s in the language under study, a matching sentence (to replace “ p ”) that, in some way yet to be made clear, “gives the meaning” of s . One obvious candidate is just s itself, if the object language is contained in the metalanguage; otherwise a translation of s in the metalanguage. As a final bold step, let us try treating the position occupied by “ p ” extensionally: to implement this, sweep away the obscure “means that,” provide the sentence that replaces “ p ”

with a proper sentential connective, and supply the description that replaces “*s*” with its own predicate. The plausible result is

(T) *s* is T if and only if *p*

What we require of a theory of meaning for a language L is that without appeal to any (further) semantical notions it place enough restrictions on the predicate “is T” to entail all sentences got from schema T when “*s*” is replaced by a structural description of a sentence of L and “*p*” by that sentence. (Davidson 2001c, p. 23)

Davidson notes that the condition placed on (T) “is in essence Tarski’s Convention T that tests the adequacy of a formal semantical definition of truth” (ibid.), and so the result of focusing on getting the right relation between the sentence *s* and the sentence that goes in the place of “*p*” in “*s* means that *p*,” but with an extensional “filling,” is to require that we have a truth theory for the language that meets Tarski’s requirement of adequacy.

Davidson, as can be observed in the above passage, immediately goes on to suggest that one can put constraints on a truth theory that will ensure that it meets Convention T (or a suitable analog for a natural language) without appeal to any semantic concepts apart from concepts drawn from the theory of reference (truth, satisfaction, and reference). As we read Davidson, this is an extension of the project of providing a constructive account of natural languages, which looks beyond illuminating the relation between semantical primitives and complexes to showing something about how the concept of meaning is related to other concepts. On this account, Davidson hoped initially to show that an axiomatic truth theory adapted to a language with context sensitive elements would satisfy an appropriate analog of Convention T if it were merely true, because of the added sensitivity required for the theory to predict correctly the truth of sentences containing demonstratives, among other context sensitive elements, in such sentences as “This is grass,” “This is snow,” “This is white,” and “This is green.” This would rule out, for example, such spurious T-sentences as: “Snow is white” is true for a speaker at *t* iff grass is green at *t*.

In (Lepore and Ludwig 2005, chs. 4, 5, and 9) we show that the initial and the extended project can be separated, and that it suffices to carry out the initial project to have a certain body of knowledge about a truth theory, roughly:

- (i) that its axioms use metalanguage expressions in giving satisfaction and truth conditions that interpret the object language terms for which they are used to give satisfaction conditions;
- (ii) what the axioms as given by the theory mean;
- (iii) a grammar for the language of the truth theory to be used in connection with
- (iv) a canonical proof procedure that draws only on the content of the axioms in proving T-sentences.

If the axioms meet condition (i), we will call them and the theory they determine interpretive. (i)–(iv) do not involve the assertion of the axioms of the truth theory, and the language in which they are stated need not be the language of the truth theory. They are supposed to state knowledge that anyone could have about a truth theory that would suffice to use it to interpret object language sentences. This does require being in a position to understand the truth theory. (ii) and (iii) are supposed to secure this: knowing what the axioms mean and having a grammar for the language of the truth theory, we are in a position to come to know the language of the truth theory. Given this: if one knows (i)–(iv) about a truth theory, one is in position to derive a T-sentence, “*s* is true iff *p*” (ignoring context sensitivity for the sake of simplifying exposition) for any object language sentence *s*, which draws only on the content of the axioms. Knowing this, we know that “*p*” in the theorem translates *s*. Knowing this, we know that the corresponding instance of (M) is true. Understanding the language of the truth theory, we are then able to infer from (M) being true to its being the case that *s* means that *p*.

The initial project is concerned just with explaining how understanding semantically primitive expressions (those not understood on the basis of any structure involving contained expressions which are understood independently) and their modes of combination determine how complex expressions are to be understood. Thus, it aims to reveal semantic structure in a language.

Davidson saw revealing semantic structure as of a piece with revealing logical form or grammar, and he emphasized the importance of separating this project from the analysis of lexical meaning or of concepts (Davidson 2001c, p. 31). An “adequate account of the logical form of a sentence,” according to Davidson,

must lead us to see the semantic character of the sentence—its truth or falsity—as owed to how it is composed, by a finite number of applications of some of a finite number of devices that suffice for the language as a whole, out of elements drawn from a finite stock (the vocabulary) that suffices for the language as a whole. To see a sentence in this light is to see it in the light of a theory for its language, a theory that gives the form of every sentence in that language. A way to provide such a theory is by recursively characterizing a truth predicate, along the lines suggested by Tarski. (Davidson 2001d, p. 94)

This introduces what has been an influential conception of logical form: namely, as that semantic structure in a sentence revealed in a compositional semantic theory for the language which abstracts away from the particular content of those of its primitive devices for which recursive axioms do not need to be given, though not from their general role (see also Davidson 2001b, pp. 105–6, 137–46).

Semantic structure is thus revealed in the recursive devices of the language by which more complicated expressions of the same category are built up out of simpler expressions, and in the role of types of simpler expressions as expressed in the rules detailing their contributions to truth conditions (for example, the role of a predicate, of one or more places, or the role of a referring term). The semantic structure of a

community's language is determined by the dispositions of its speakers. A theory of semantic structure of a language for a given community is then responsible to the competencies of its speakers in using it, and we should expect there to be a correspondence between what the theory reveals about the semantic structure of the language and the structure of the complex of competencies of speakers of it. It is in virtue of this that an adequate compositional meaning theory would be thought to show how it is possible for finite speakers of a language with infinite expressive resources to be in a position to understand any potential utterance of a sentence of the language. In the case of the use of an interpretive truth theory in giving a compositional meaning theory, the idea is that the axioms attaching to primitive terms represent rules for their use that are expressed in corresponding competencies of the speakers of the language, in the sense that they use the words in accordance with the rule. Thus, for example, if an axiom says that a predicate is satisfied by something iff it is red, then a speaker will use it in a sentence where it commits him to a thing's being red only if the speaker is confident that it is red (on the assumption that he is being sincere and not engaged in figurative speech, acting in a play, horsing around, and so on).

The projects of revealing logical form and capturing semantic competence are therefore interconnected. To show what the logical form of a sentence in a natural language is we seek to reveal its truth-relevant semantic structure. To show its truth-relevant semantic structure is at the same time to express the rules for the use of types of expressions (in combination with others of their type) that mirror (in the sense above) the semantic competence of speakers of the language.

The application of a formal theory of truth to a natural language requires some regimentation. First, structural and lexical ambiguities must be removed. Second, for the application of the usual clauses for quantifiers, explicit variables must be introduced. Third, some regimentation may be required because it requires making explicit semantic features of sentences that are not represented syntactically. For example, if Davidson's account of the logical form of action sentences is correct (Davidson 2001b), then [1] involves an implicit quantifier over events and the adverbs contribute predicates of them, as illustrated in [2], which would be said to give the logical form of [1].

[1] Lincoln spoke quietly at Gettysburg, with a pronounced Kentucky accent.

[2] There is an event e such that e was a speaking by Lincoln and e was quiet and e occurred at Gettysburg and e was produced with a pronounced Kentucky accent.

The truth theory is applied to the regimented portion of the language, and to the rest of the language through a mapping of sentences, in use in order to disambiguate, if necessary, to sentences of the regimented language. The account given by a theory of the logical form of the sentence is given by its regimentation, if any, and the canonical proof of the T-theorem for it.

Davidson held that the logical form of a sentence was relative to the deductive theory of the truth theory, to how we choose the regimented portion of the language,

and to how we then regiment sentences in the unregimented portion. Appeal to semantic competence can help to constrain the latter. We have suggested that the relativity to a deductive theory can be overcome by thinking of sameness of logical form between two sentences ϕ and ψ as requiring that there be interpretive truth theories for the language of each in which there are corresponding canonical proofs of interpretive T-theorems for ϕ and ψ —proofs which reveal the same semantic structure (Lepore and Ludwig 2002; 2007, ch. 13).

Vagueness and the semantic paradoxes present a special challenge to truth-theoretic semantics. Suppose that B is a borderline case for “bald”; that is, someone whose head our practices with “bald” seems to leave unclassifiable because it is somewhere between a case where we feel confident that using “bald” is licensed and a case where we feel confident that to predicate “bald” of it would be violating the rules for its use. “B is bald” would therefore seem to be neither true or nor false, and hence the T-sentence for it (TB) would likewise seem neither true nor false.

(TB) “B is bald” is true iff B is bald.

No truth theory that issues in such a theorem could be true. Similarly, if we allow that natural languages like English are universal, then it seems that we ought to be able to name any sentence and to predicate truth or falsity of it. Thus, it would seem possible to name a sentence that says of itself that it is not true, as in (L):

(L) L is not true.

The T-sentence for this will be (TL), which is necessarily false, and which by substitution of “L” for “L is not true” yields a truth-functional contradiction in (TL★):

(TL) “L is not true” is true iff L is not true.

(TL★) L is true iff L is not true.

Davidson made two basic suggestions about the semantic paradoxes (Davidson 2001c, pp. 28–9). The first was that not much would be lost by just giving a truth theory for that portion of the language that does not include semantic predicates. The second was that we might deny that these contradictions could be derived in natural languages on the grounds that the semantic predicates in natural languages could be plausibly construed as expressing ambiguously a hierarchy of truth predicates, each applying only to sentences of a lower order (Lepore and Ludwig 2005, pp. 135–6). For the case of vagueness, Davidson suggested that we aim to prove T-sentences that use vague metalanguage predicates that match the use profiles of the object language predicates whose contributions to truth conditions they are used to express. This does not, however, secure the truth of the truth theory, even if it still allows one in a certain sense to read off the meaning of the object language sentence from the T-sentence. An alternative response suggested by Lepore and Ludwig (2005, pp. 140–1) is to take seriously the idea that the meaning theory is distinct from the truth theory. If we take the meaning theory to be the body of knowledge that one has to have in order to

interpret any potential utterance of an object language sentence on the basis of understanding its semantical primitives and their mode of combination, then this consists basically of the knowledge specified in (i)–(iv) above (for further discussion see Lepore and Ludwig 2005, ch. 9; 2011). Although this puts one in a position to state an interpretive truth theory for the object language, it does not require that one assert its axioms. Consequently, commitment to the meaning theory does not involve commitment to the truth theory. Of course, we want to derive M-theorems, but the M-theorems that correspond to (TB) and (TL)—namely, (MB) and (ML)—are true, not false.

(MB) “B is bald” means that B is bald.

(ML) “L is not true” means that L is not true.

Distinguishing in this way between the meaning theory and the truth theory, then, allows us to provide meaning theorems for semantically defective sentences in a straightforward way.

2 The extended project and radical interpretation

As we noted above, at the point at which Davidson proposed that a truth theory could be exploited to achieve the aims of a compositional meaning theory, he also proposed that deeper insight into meaning could be achieved by placing non-semantic constraints on it that ensure that it meets an analog of Convention T for context-sensitive languages. In “Truth and Meaning” he suggested that the additional resolving power of context sensitivity in natural languages suffices for any true truth theory to satisfy an appropriate analog of Convention T for natural languages. He later recognized that this was not adequate because, given any extensionally adequate predicate axiom, one can construct another one which is true iff the first is, though the metalanguage verb phrase used to give truth conditions is not the same in meaning. For example, one can add a conjunct to the satisfaction conditions for any predicate without change of truth-value if it is necessarily true. Thus, to any predicate axiom’s satisfaction conditions we can add “and $2 + 2 = 4$ ” without change of truth-value. One true theory might then issue in the T-theorem “‘Mars is red’ is true iff Mars is red,” and another in the T-theorem “‘Mars is red’ is true iff Mars is red and $2 + 2 = 4$.” Both theories are true, but only one satisfies Convention T, and context sensitivity clearly does not help here. This example shows also that it is no help to require that the axioms be laws or that they be necessary or analytic. Call this the extensionality problem.

Davidson had thought all along of the truth theory as an empirical theory for a speaker’s or a community’s language. His response to the extensionality problem was to return to this thought and to suggest that a truth theory would be adequate for interpretation if it were, or would be, a theory that a radical interpreter would find satisfactory in the light of behavioral responses to the environment and to other speakers. What is a radical interpreter? The standpoint of the *radical* interpreter

excludes, ultimately, all evidence other than purely behavioral evidence, and in particular excludes any description of the subject of interpretation that brings to bear intentional or semantic vocabulary. Davidson took this standpoint to be the fundamental one from which to investigate meaning and propositional attitudes on the grounds that speech is by its nature a social phenomenon:

Mental phenomena in general may or may not be private, but the correct interpretation of one person's speech by another must *in principle* be possible... [W]hat has to do with correct interpretation, meaning, and truth conditions is necessarily based on available evidence. That meanings are decipherable is not a matter of luck; public availability is a constitutive aspect of language. (Davidson 1990, p. 314)

Restricting adequate theories to those that can be confirmed from the standpoint of the radical interpreter provides a constraint on a truth theory, motivated by the nature of its subject matter, that can serve the goal of illuminating meaning rather than presupposing it.

As a practical matter, Davidson helped himself to “an intermediate stage” (Davidson 2001f, p. 161) in theory construction: namely, to hold-true attitudes. A hold-true attitude is a belief to the effect that a sentence is true. Hold-true attitudes are, as Davidson puts it, vectors of meaning and belief: if one believes that p , and knows that s means that p , then one will believe or be disposed to believe that s is true. (We will accept for the sake of argument the assumption that Davidson makes that for every belief a speaker has, he holds true a sentence that expresses it; this may be made more plausible by thinking of hold-true attitudes as related to dispositions to sincere assertion of sentences as appropriate.) Davidson reasoned that if we could identify hold-true attitudes, and then hold one of the two factors that determine them fixed, we could solve for the other. Since we aim to uncover meanings, Davidson proposed that we hold fixed belief by assuming that speakers are mostly right, and in particular that they were mostly right about their environments. Given this, we can correlate hold-true attitudes that are responses to the environment with what they are responses to, and treat the conditions they are responses to as what the belief is about which is the basis for holding the sentence true, in light of its meaning. This assumption, together with the assumption that the speaker is, as an agent, constitutively largely rational, Davidson called the Principle of Charity. Later he distinguished the two components as the Principle of Correspondence and the Principle of Coherence (Davidson 1985, p. 92; 2001f, p. 211). With the Principle of Correspondence in hand, we first gather evidence to arrive at generalizations of the form (L). Then on the assumption that we have identified a condition, in the use of “ p ,” which the belief on the basis of which the speaker S holds true s at t is about, we tentatively conclude that the corresponding instance of (T) is interpretive.

- (L) S holds true s at t iff p .
- (T) s is true for S at t iff p .

We assemble by this method a set of target T-theorems for an interpretive truth theory for the speaker's language, and then aim to construct a truth theory that has them as canonical theorems. As we build up a picture of the speaker's language, we must at the same time build up a picture of the speaker as an agent, and so fill in our picture of the speaker's propositional attitudes, not just his beliefs but also his desires, interests, intentions, plans, and so on. We test our evolving theory against new observations and make adjustments to improve the fit of theory and evidence, which may involve the rejection of some of the initial target T-theorems in favor of the hypothesis that the speaker has been involved in some sort of systematic mistake.

3 Indeterminacy of interpretation

Davidson held that, after all the evidence was in, there would still be different and equally acceptable interpretation theories—theories which assign different truth conditions to a speaker's sentences—even different truth conditions to the same sentence that yield different verdicts on its truth-value. The differences would be due to differences in starting points and differences in decisions about when to attribute error or irrationality and how much error and irrationality to attribute to the subject. On the face of it, this seems to show that the radical interpreter's evidential base is not adequate, by his own lights, for confirming an interpretation theory, because it shows that the evidence underdetermines the correct theory. However, given his commitment to taking the standpoint of the radical interpreter as basic, Davidson takes this to show not that the correct interpretation theory is underdetermined by the evidence, but instead that all empirically adequate theories capture equally well all the relevant facts about meaning. We say, then, that there is indeterminacy in interpretation in the sense that there are intuitively incompatible interpretation theories that capture all of the meaning facts equally well. What is to make sense of this is the idea that the concepts deployed in an interpretation theory have their content exhausted by how they help us to organize the evidence that is brought to bear on confirming the theory. We may call such concepts "theoretical concepts."

There is, of course, an air of paradox about this. How can a theory T which, for example, interprets "Bugs est un Lapin" as meaning (univocally) that Bugs is a rabbit, and a theory T' which interprets "Bugs est un Lapin" as meaning (univocally) that Bugs is a squirrel, both be correct? Davidson sought to remove the air of paradox by comparing different interpretation schemes with different schemes of measurement: that is, different ways of using numbers to keep track of physical quantities, such as the Fahrenheit and Centigrade scales. There are different systematic ways of mapping the numbers onto physical quantities, because the structure of relations in the domain to which numbers are mapped is not as rich as the structure of relations among the numbers. For example, in the physical measurement of temperature—say by the height of a column of mercury in a glass tube—we can make sense of two differences in temperature being the same: that is, sameness of interval. We keep track of this by

assigning to each the same interval in our measurement system. But the absolute difference between the numbers we use cannot be correlated with any measurement outcome. So we are free to assign an interval as measured as the difference between 0 and 100 or as the difference between 32 and 212. What this analogy suggests about how to understand the position of the radical interpreter is that in mapping his sentences onto those of a speaker, he finds fewer distinctions marked in the speaker's language than in the language he brings to bear in interpreting it. In consequence, he finds that there are a number of different systematic ways of mapping his sentences on to those of the speaker that capture equally well the facts of the matter. Each interpretation scheme is like a measurement scale, and one can in principle see, relative to the different initial decisions that led to the divergence, how each keeps track of the same information.

Does the analogy remove the air of paradox? It is not clear that it succeeds, for it presupposes that the interpreter can make distinctions that are not present in the speaker's language. It must be possible for another to speak the interpreter's language. But this could not emerge from the radical interpreter's standpoint. It therefore appears that we must admit after all that the evidence for interpretation from the standpoint of radical interpretation is not sufficient to interpret every speaker (for further discussion see Lepore and Ludwig 2005, ch. 15).

4 First-person knowledge and authority

This brings us back to the question of whether the concepts that the interpreter deploys in interpretation are in the sense above theoretical concepts. If they are, then their content should be exhausted by their use in organizing the evidence on the basis of which a theory deploying them is confirmed. We noticed that from the standpoint of the interpreter himself, it appeared that there were distinctions to be drawn among theories that the evidence available cannot discriminate. This presupposes that the interpreter has access to those distinctions in some way other than interpretation. This distinct perspective on meaning, and thought as well, is an independent threat to the claim that the concepts of the theory of interpretation are theoretical concepts. For if the conceptually fundamental basis on which to apply these concepts is the third person point of view, it is *prima facie* puzzling that there should be what appears to be a completely independent perspective—the first person perspective—from which to deploy them. Minimally, what is required to reconcile these is a way of explaining the first-person perspective from the standpoint of the third-person perspective, and particularly that of the radical interpreter.

Davidson took up this challenge primarily in “First Person Authority” (Davidson 2001j). The argument of “First Person Authority” focuses on the asymmetry between the authority of a speaker to say what he thinks and that of an interpreter. Davidson argues that the interpreter has no choice but to assume that the speaker knows the meanings of his words, while there need be no such presumption that the interpreter

does. Then, given equal knowledge of what sentences the speaker holds true, desires true, and so on, there will be a presumption that the speaker knows what he thinks, but no presumption that the interpreter does. In Lepore and Ludwig (2005, ch. 20), we argued that this argument has the order of explanation backwards, because hold-true attitudes are supposed to be the result of a speaker's knowing what he believes and knowing that what he believes is expressed by a sentence of his. In addition, the explanandum appears to be too narrow. It is not merely that any subject is to be presumed to know what he means while an interpreter is not, but also that the subject is presumed also to know what he thinks, and what he believes generally, including what he believes about what sentences of his are true.

There appears, however, to be a shorter route to the conclusion that Davidson wants than the one he takes: namely, that it emerges from the standpoint of radical interpretation that the speaker is to be presumed to know both what he means by his words and what he thinks. For the interpreter's procedure, as Davidson describes it, requires identifying hold-true attitudes and correlating at least a subset of them with varying conditions in the speaker's environment. The Principle of Correspondence advises taking the speaker's beliefs about his environment to be true. But this gives us a reason to think that the condition correlated with the hold-true sentence can be treated as giving its interpretive truth conditions only if we assume that the speaker in coming to hold true the sentence did so on the basis of knowing what it means and knowing that it expresses what the speaker believes about the environment on the occasion—which requires the speaker to know what he believes. It is difficult, in fact, to see how any procedure for interpretation could avoid the assumption that speakers know what they believe, because speech is produced on the basis of what speakers believe about their environments and think their sentences mean. We are thereby able, for a certain class of utterances, to let the conditions under which the utterances are made to serve as a guide to what the sentences mean. But this requires us to think that the speaker chooses sentences to utter that reflect what the speaker believes, given what they mean, which implies that the speaker knows what he believes.

However, establishing merely this—namely, that interpretation can succeed only if the speaker knows what he thinks and means—does not shed much light on the question of how one applies the concepts in the first-person case. For this reason, it falls short of what is needed, which is to show that the application of these concepts from this perspective does not show that they have a life independent of their deployment from the third-person point of view.

5 Our knowledge of the external world and externalism about thought content

In “Empirical Content” (Davidson 2001h) and “The Coherence Theory of Truth and Knowledge” (Davidson 2001i), Davidson began to draw out some of the consequences

of taking the position of the radical interpreter to be the fundamental standpoint from which to investigate thought and meaning for our knowledge of the external world and the relation of thought to our environments.

Traditional skepticism about the external world rests on two assumptions. The first is the logical independence assumption: that the domain of the mental about which we have non-inferential knowledge is logically independent of the external world. The second is the epistemic priority assumption: that facts about the mental (in this sense) are epistemically prior to facts about the external world, in the sense that any justification for a proposition about the external world must rest on an argument that starts with facts about the mental. Given the second, we must justify beliefs about the external world on the basis of facts about our mental lives. Given the first, we cannot do so because it precludes any *a priori* route, and the *a posteriori* route requires independent access to the external world.

Davidson took aim at the first of these two assumptions. As we have noted, he took the standpoint of the radical interpreter to be the conceptually fundamental position from which to investigate meaning and thought. This was motivated by the thought that evidence for meaning had to be intersubjectively available because language is a social phenomenon. Davidson took this to mean that all the relevant evidence was behavioral, and that the expression of a language in behavior (taken in a broad sense of causal interaction with one's environment and others) was necessary to having a language, and that consequently any speaker would in fact be interpretable in principle on the basis of that evidence. When we combine this thought with the observation that the application of the Principle of Correspondence is necessary in interpretation, we immediately come to the conclusion that linguistic beings must have mostly true beliefs about their environments. Davidson independently argued that having a language is necessary for thought (Davidson 1975). When these two observations are put together, we obtain the view that, as Davidson puts it, "belief is in its nature veridical" (Davidson 2001i, p. 146). This provides a transcendental guarantee (as a condition on the possibility of having any thoughts at all) that our picture of the world is by and large correct. And this entails the falsity of the logical independence assumption.

It follows likewise from taking the Principle of Correspondence to express a constitutive principle governing belief that beliefs are relationally individuated. For the Principle of Correspondence requires that by and large environmentally prompted beliefs are about the conditions that prompt them. Since the contents of beliefs are coordinated with the contents of other attitudes in rational patterns in the production of behavior, it follows that the propositional attitudes more generally are relationally individuated. If we were to hold fixed all the non-relational facts about a speaker but vary the causes of beliefs, given that they are governed by the Principle of Correspondence, it would follow that their contents would vary as well. Thus, externalism about thought content is a consequence of taking the standpoint of the radical interpreter to be conceptually fundamental in investigating meaning and thought.

Given a transcendental guarantee that most of our beliefs, especially our environmentally directed beliefs, are true, Davidson argued that paying attention to how well a given belief hangs together with one's general picture of the world is a guide to the likelihood of its truth. That is to say, we can be assured that coherence of a belief with other beliefs makes for correspondence. Thus we can justify our beliefs by appeal to how well they fit into a coherent body of belief about how the world around us is.

Davidson thought the mistake of the tradition was, in part, to seek to justify beliefs by appeal to sense data, percepts, or some other such epistemic intermediary. The mistake, he thought, lay not just in the assumption of the logical independence of the nature of such epistemic intermediaries from the nature of the external world, but in even thinking that percepts, or sense data, or the uninterpreted given, could in principle provide a reason for belief. The trouble, as Davidson saw it, was that, in so far as these were taken to be uninterpreted, they could not provide any reason for belief, because reasons must involve representations of how things are. Thus, he concluded, only beliefs can provide reasons for beliefs. If we are giving reasons for one or other particular belief about the external world, we can cite others. If we are trying to provide a reason for trusting our beliefs about the external world in general, however, we cannot cite other beliefs about the external world because their status as justificatory is in part what is in question.

Davidson concluded that the challenge is to "find a *reason* for supposing most of our beliefs are true that is not a form of *evidence*" (Davidson 2001i, p. 146). The reason he then finds lies in the connection between meaning and thought, on the one hand, and meaning and true belief, on the other. In the resulting picture of how we make epistemic contact with the world, experience drops out of the picture altogether. It is a causal intermediary between us and the world around us, but on Davidson's view it plays no epistemic role whatsoever.

6 Triangulation

The idea of the triangulation of two speakers interacting with each other in response to a common object was introduced in Davidson's work in the 1990s in response to a problem that he had not responded to explicitly becoming salient for him. We first discuss the problem, and then assess how the idea of triangulation is supposed to provide a response.

The Principle of Correspondence states that a speaker's beliefs about his environment are by and large true and about what prompts them. This encourages us to think that, by identifying the systematic causes of a speaker's hold-true attitudes, we identify what the contents of the beliefs are on the basis of which the speaker holds true those sentences. Even if we grant this, however, there is a problem, which is connected with the charge we reviewed earlier that the evidence that the radical interpreter has available underdetermines the correct theory. The difficulty is that there are generally many systematic causes of our beliefs, because causal influence is transmitted through

causal chains. It is not plausible to say that it does not matter on which of them we focus. For example, some pattern of stimulus of the retinas for each person is correlated, *ceteris paribus*, with his holding true “There is a television set in front of me,” and is in turn reliably correlated, *ceteris paribus*, with there being a television set in front of him. It seems implausible to say that there is no fact of the matter whether the speaker is talking about something in his environment (a television) or about the stimulation of his sensory surfaces. But what in the constraints on the radical interpreter’s procedure prevents us from interpreting people as talking mostly about what is going on at their sensory surfaces rather than in the environment around them?

Triangulation is Davidson’s answer to this question. The motivation for the restriction on the radical interpreter’s evidence is that language is a tool for communication. Thus we should think of interpretation as taking into account not just a speaker in isolation from other speakers but in interaction with them. The smallest unit of conversation consists in two speakers. When they communicate about something in their environment, there must be a common cause of their responses to it for them to be in communication. Their responses to each other, and to the common cause of their responses to the environment, form a triangle. In interpretation, we seek to find, therefore, the common cause of their responses, and we identify that as what their beliefs are about. In this way, we seek to find the link in the causal chains leading to each of their beliefs that gives them a common content. This does not mean that the radical interpreter has to find two speakers he can observe in order to interpret them, however, for the interpreter himself can play the role of the second speaker. In practice, this comes to his finding the other speaker to be thinking and talking about things that the interpreter finds salient.

Does triangulation solve the problem of locating the right causes of beliefs to identify as giving their contents? Thinking of it from the point of view of the interpreter finding salient causes of a speaker’s hold-true attitudes, which guarantees that he has settled on a common cause, may make it seem as if it has to be right. But it is not clear that the interpreter is entitled to his assumption that he has hit upon the cause that gives the content of the speaker’s beliefs if there are other common causes of the environmentally directed beliefs of both which the interpreter does not typically think about. It is clearly possible for different creatures to find different things in the environment salient. Of course, the interpreter may help himself to the assumption that the speaker whom he is interpreting is likely to find the same things salient. But this is just to assume that the speaker is thinking for the most part about the same things. It is not to identify objective structures in the speaker’s interaction with the environment (which includes other speakers) that determine what the speaker is thinking about. To rely on knowing what one is thinking about oneself and the assumption that any potential speaker would be thinking in the same environment about more or less the same things is to give up on the idea that it can be read off from objective features of the environment: that is, that it can be constructed wholly from the third-person point of view. It must be, then, that there is in fact only one thing that can play the role of the

common cause of two speaker's responses to the environment if triangulation is to solve the problem.

But it is not at all clear that this is so. The same problem in fact reappears when we add an additional speaker to the picture. Consider two people sitting on a couch in front of a television that is tuned to a news broadcast. Suppose that the announcer has just said "The mayor of Kandahar was assassinated in southern Afghanistan last night." Suppose as a result each speaker has come to hold true the sentence, "The mayor of Kandahar has been assassinated." The proximal external causes of their hold-true attitudes are the pattern of vibrations on their eardrums, and so not a common cause. But those distinct causes have a common cause in the pressure waves in the air propagated from the television's speakers, which are caused there by the vibrations of the speaker's diaphragms, which are caused by electrical signals generated by the television, which through a series of links through radio waves or signals on a coaxial cable can be traced back to the television station, to a satellite in orbit, and eventually to events in Kandahar, and beyond. Which of these common causes is the one that their thoughts are about? To say it is the one that each would find the natural one on which to focus in interpreting the other is just to say that it is the one that they are thinking about, but not to give any objective criterion for picking it out.

This completes our brief overview of the complex of issues in Davidson's philosophy that forms the backdrop for the essays in this volume. We turn now to an even briefer overview of the contributors' essays.

7 Contributions to the volume

Part I of the volume comprises six essays on themes connected with Davidson's work on the theory of meaning and logical form. Part II comprises six essays that are concerned with the possibility of radical interpretation and its implications for our understanding of the mind and our epistemic relations to ourselves and to the world. In the following we provide a brief and far from exhaustive sketch of the goal of each, sometimes selecting a few points for more discussion, mostly where they bear directly on our interpretation of Davidson. In this we will still fall far short of doing justice to the interest and intricacies of these essays.

Part I

"Davidson's contribution to the philosophy of language"

Gilbert Harman

Gilbert Harman provides a lucid overview of Davidson's contributions to the philosophy of language. We agree with Harman that Davidson's emphasis on compositionality and the value of exploiting a truth theory in capturing semantic competence are among Davidson's chief contributions to the philosophy of language. Harman accepts also that there are two projects in Davidson's philosophy of language corresponding to what we called above the initial and extended projects.

In surveying the development of the field and thus Davidson's influence, Harman notes that Davidson's demand of compositionality, and insistence on the importance of showing how to integrate accounts of semantic structure into truth theories for a language, has won the day, but also that Davidson's focus on first-order extensional theories has not. And though most philosophers and linguists endorse Davidson's treatment of adverbs and verbs of action, his treatment of attitudinal attributions has been far less influential.

Harman argues that Davidson's use of truth theories in pursuit of giving a compositional meaning theory for a natural language does not illuminate lexical meaning for non-logical particles. We agree with this point, and we believe that Davidson would have agreed as well. This has to do with the goals of what we called the initial project. However, we also hold that Davidson's extended project, namely, treatment of the truth theory as an empirical theory, and the adequacy constraint he places on it of being confirmable from the standpoint of a radical interpreter, was supposed to help provide illumination of more than just the semantic structure of the language.

The charge that Davidson does not, and does not intend to, illuminate the meanings of primitive expressions is the charge, as Dummett once put it, that Davidson's theory of meaning is modest in the sense that it does not seek to explain the concepts expressed by its primitive vocabulary, but to convey meaning only to someone who already grasps the concepts expressed by the primitive vocabulary (Dummett 1975). While Dummett himself once leveled this charge against Davidson, in later work he rejected it. He puts it this way, in one place: "The reason why a meaning-theory of Davidson's kind is not after all a modest one is that, contrary to the way he presents it, the so-called evidence is not an external support on which we rest our confidence in it but, rather, is integral to it; it is part of the theory itself" (Dummett 1991, p. 109). As we would put it, the idea is to illuminate what it is for words (in general) to mean what they do by seeing how public evidence available to speakers that does not presuppose any knowledge of meaning can be marshaled in support of an interpretive truth theory; this project must relate a speaker's use of words to his environment through his use of them in his sentences, and maximize intelligibility in light of the totality of evidential and *a priori* constraints. If we had this—that is, a detailed account of how to do this for a particular language—then we would have insight not just into semantic structure in the language, but insight into what it was for its primitive expressions to mean what they do as well.

Harman also nicely charts out similarities and differences between Davidson and Quine on radical interpretation/translation; and where they parted ways on the significance of indeterminacy with respect to translation, interpretation, and reference. Finally, Harman usefully draws attention to important differences between Quine and Davidson on whether we can make sense of others having a language without being able to translate it into our own: whereas Quine held we could, Davidson held famously that we could not.

“Truth theories, competence, and semantic computation”*Peter Pagin*

Peter Pagin takes up two main issues. The first is the question of in what sense Davidson intended an adequate truth theory for a natural language to explain how it is possible for speakers of it to understand any potential utterance of a sentence of the language on the basis of what its semantically primitive components mean and their mode of combination. The second is whether, in that sense, it is plausible that a truth theory can do the job.

Pagin formulates the explanatory requirement as follows.

(ER) An adequate T-theory for a language L explains how it is possible for speakers of L to effectively determine the meaning of any meaningful expression of L. (this volume, p. 49)

The reason why a question arises about what Davidson had in mind is that there is some reason also to attribute to him a second view on the adequacy of a truth theory: namely, that only its theorems matter to its interpretational adequacy—a claim that Pagin formulates as (NPR).

(NPR) Only the *theorems* of a T-theory are relevant to its interpretational adequacy. (ibid., p. 52)

In terms of the discussion above, this is the view that it does not matter whether the theorems of the truth theory are interpretive, as long as the theory satisfies Convention T. Pagin puts it this way: “The actual cognitive architecture need not correspond at all to the structure of the T-theory. The structure is needed for generating the theorems but only the theorems generated are relevant to its interpretational adequacy” (ibid.).

In support of this view, Pagin cites two passages from “A Nice Derangement of Epitaphs” (Davidson 2005a) and a passage from “Radical Interpretation” (Davidson 2001e). The passages in the former establish that Davidson did not intend to be ascribing propositional knowledge of truth theories to speakers to explain their semantic competence, nor to be requiring that an adequate truth theory provide any guide to the “details of the inner workings of some part of the brain” (Davidson 2005b, p. 96). It is not clear that they establish that the cognitive architecture need not correspond at all to the structure of the T-theory, at least if this is read as meaning that the theory is not to model in any way semantic competence beyond tracking correct output. When Davidson says, “Here however there is no reason to be concerned with the details of the theory that can adequately model the ability of an interpreter” (ibid., p. 95), he is not saying that it is not one of his goals, but that in the context of the main thesis of “Nice Derangements” this can be put aside. For he follows up by saying: “All that matters in the present discussion is that the theory has a finite base and is recursive, and these are features on which most philosophers and linguistics agree” (ibid.). He goes on

to say that the claims he advances about “what would constitute a satisfactory theory are not . . . about the propositional knowledge of an interpreter” or about “the inner workings . . . of the brain,” but “rather . . . about what must be said to give a satisfactory description of the competence of the interpreter” (ibid., p. 96). A satisfactory description of *linguistic competence*, however, has to take into account not just sentence-level competence but also competencies attaching to the rules for the use of semantically primitive vocabulary items. A number of commentators on Davidson, including the authors of this Introduction, have taken this to be a commitment to seeing the axioms of the theory corresponding to dispositions of speakers to use words, in the sense discussed above, which is compatible with remaining silent on what realizes the dispositions of the speaker (Lepore and Ludwig 2005, pp. 31–3, 121–4).

Pagin also cites this well-known passage from the beginning of “Radical Interpretation”:

What could we know that would enable us to [interpret another’s speech]. How could we come to know it? The first of these questions is not the same as the question what we *do* know that enables us to interpret the words of others. For there may easily be something we could know and don’t, knowledge of which would suffice for interpretation, while on the other hand it is not altogether obvious that there is anything we actually know which plays an essential role in interpretation. (Davidson 2001e, p. 127)

This seems largely silent on the question of whether the structure of the truth theory Davidson eventually proposes to use should aim to capture something about the structure of the competence of speakers for which it is a theory. For the knowledge in question here is propositional knowledge, and it is compatible with denying that speakers of a language have propositional knowledge of a truth theory to see the truth theory as aiming to capture the structure of a complex practical ability. Pagin, however, also makes the point that in radical interpretation the interpreter gains access to a speaker’s language by the assumption of Charity, which secures target theorems of a truth theory that is to meet Convention T (or its analog for natural languages), and does not impose directly any requirement on the axioms. While this is true, it leaves open the question of what the constraint Davidson imposes is *supposed* to secure, and, in particular, whether it is supposed to secure that the axioms are in fact interpretive.

These interpretive points bear on what explanatory requirement Davidson thought truth theories ought to meet, and so on what it is necessary to show in order to show that a truth theory meets that requirement. Pagin suggests that the way to see how [ER] and [NPR] are compatible is to take [ER] not to be asking for an explanation of how actual speakers are able to understand the languages they speak, but rather to be asking for an explanation of how the languages actual speakers speak could be understood: that is, some mechanism or other that could be instantiated by a speaker of the language (with the same general limitations as human speakers—otherwise the question becomes of little interest) that would enable them to understand any expres-

sion of the language. On this view, the truth theory (together with some knowledge about it) is introduced as one way that a speaker of the language could understand it.

With this in mind, Pagin takes up the question of whether appropriate knowledge of a truth theory does suffice to explain how a speaker of a natural language could understand it. Commentators have just assumed this. Pagin asks what it would take to show that this assumption is not misplaced. Pagin suggests reformulating the question as a question about whether a truth theory provides a time-efficient method of determining the meanings of sentences of a natural language. For if a truth theory provides an in-principle way of determining meanings but one which is impractical for human beings, it would not serve in any interesting sense as an explanation of how it is possible for us. Pagin suggests that if a truth theory provides a method for determining meanings that can be computed in polynomial time, it can explain how human speakers could understand their languages. If they can be used to determine meanings only in exponential time, however, they do not provide an adequate explanation of how it is possible. As an encouraging beginning of the task, Pagin shows in detail that for a simple first-order context-free language, "T-theories represent the problem of semantic interpretation as tractable, in fact as having low complexity, even if not minimal" (this volume, p. 70). This result is as relevant to those who think Davidson intended truth theories to capture the full structure of speaker's semantic competencies as to those who think his explanatory interests were less ambitious, for if a truth theory did not represent at least one plausible model for how the languages we speak could be understood, they could not be a model for how we actually understand them.

"Davidson's explication of meaning"

Gary Ebbs

Gary Ebbs takes issue with the interpretation of Davidson that we advance in (Lepore and Ludwig 2005). We see Davidson as pursuing the traditional project of illuminating meaning by indirection, and as rejecting not the intelligibility of meaning but instead just the utility of meanings, construed as entities, that are assigned to expressions in pursuit of providing a compositional meaning theory for a language. The key claim in our interpretation is that Davidson intended to describe in non-semantic terms constraints on a truth theory that suffice for it to meet what Ebbs calls the *s*-means-that-*p* requirement, which is essentially that the theory meet an analog for Convention T for natural languages. Ebbs argues, against our interpretation, that "Davidson regards traditional or common-sense ideas of meaning as obscure and proposes that we replace them by a notion of truth conditions characterized holistically in terms of an empirically testable truth theory for a given speaker's language" (this volume, p. 76). He argues that it is more natural to see Davidson in the historical context as undertaking an explication of meaning in Carnap's sense, and that a review of passages in Davidson's work does not tell decisively in favor of our interpretation, so that the better interpretation overall is the one he advances. The main issue here is not whether Davidson was

proposing that an empirically confirmed truth theory could be used for interpretation, but whether he aimed to shed light on the ordinary notion of linguistic communication in making this proposal. Ebbs agrees that if Davidson was aiming to put constraints on a theory that enabled it to meet the s-means-that-p requirement, then our interpretation would be correct.

It would not be appropriate to enter here into a full-scale defense of our interpretation against Ebbs, but it may be helpful to draw attention to some points we would wish to make in a longer discussion.

First, with respect to the considerations of historical context, it is not clear that Carnap's project loomed large in Davidson's thinking. Davidson trained in classical philosophy, and wrote a dissertation on Plato's *Philebus*. His interest in philosophy of language and philosophy of action developed later and became merged in the project of radical interpretation. One of the stimulants was an invitation from J. C. C. McKinsey to collaborate with him on a paper for the Library of Living Philosophers volume on Carnap, on Carnap's method of extension and intension in *Meaning and Necessity* (Carnap 1947). Davidson took over the project when McKinsey died. In an interview, Davidson said, "I didn't know anything about Carnap when I started writing it" (Davidson 2004, pp. 249–50). There is not much evidence in this that he was deeply influenced by Carnap's philosophical project. What that work did was make salient for Davidson the problem of giving a compositional meaning theory for a natural language because it made salient the difficulties of providing an adequate semantics for belief sentences. Quine's undeniable influence on Davidson's thinking about the philosophy of language came relatively late, when in 1959 Quine came to Stanford for a term at the Center for the Study of the Behavior Sciences, with the manuscript of *Word and Object*. In the interview just cited, Davidson said, "[m]y philosophy of language didn't grow out of my relationship with Quine at all" (*ibid.*, p. 258). What Davidson drew from Quine was the importance of investigating meaning from the standpoint of the interpreter of another. What he drew from Tarski was the idea that a truth theory could provide the core of a serious theory of meaning. What is distinctive about his approach derives from the way he put these two things together. Davidson had this to say about his idea in retrospect:

I think the idea that there was a way of thinking philosophically about meaning tied to the idea of getting a serious semantic theory for as much of natural language as you could—well, I was the first person to say that, and I say it in "Truth and Meaning." There I suggested that my dream was to try to do for the semantics of natural language what Noam Chomsky was doing for the syntax of natural language. (*ibid.*, p. 259)

This does not suggest that Davidson was interested in replacing the theory of meaning with something else in a Carnapian project of explication, but that on the contrary he thought that he had hit upon a novel way of pursuing the traditional project, drawing on what he "thought was good in Quine" and what he "had found in Tarski" (*ibid.*, p. 258).

Second, there are passages which Ebbs does not discuss, in which Davidson explicitly endorses the view that his goal is to place constraints on a theory of truth that suffice for it to satisfy Convention T (or an appropriate analog for a natural language). In recounting his project, as a preliminary to responding to Foster's criticisms that he has not stated knowledge about a truth theory adequate for interpretation, Davidson says:

Since Tarski was interested in defining truth and was working with artificial languages where stipulation can replace illumination, he could take the concept of translation for granted. But in *radical* interpretation, this is just what cannot be assumed. So I have proposed instead some empirical constraints on accepting a theory of truth that can be stated without appeal to such concepts as those of meaning, translation, or synonymy, though not of course without a certain understanding of the notion of truth. By a course of reasoning, I have tried to show that if the constraints are met by a theory, then the T-sentences that flow from that theory will in fact have translations of *s* replacing "*p*" [in the T-schema—see above].

To accept this change in perspective is not to give up Convention T but to read it in a new way. Like Tarski, I want a theory that satisfies Convention T, but where he assumes the notion of translation in order to throw light on that of truth, I want to illuminate the concept of translation by assuming a partial understanding of the concept of truth. (Davidson 2001g, p. 173)

Since the *s*-means-that-*p* requirement is essentially the requirement that the truth theory meet Convention T, we may restate what Davidson says here in this way: "Like Tarski, I want a theory that satisfies [the *s*-means-that-*p* requirement], but where he assumes the notion of translation in order to throw light on that of truth, I want to illuminate the concept of translation by assuming a partial understanding of the concept of truth." The burden of the explication reading of Davidson is to explain away his attributing to himself a view that is incompatible with it.

"Against logical form"

Zoltán Gendler Szabó

Zoltán Szabó argues that there is nothing that answers to the idea of logical form as an objective feature of a sentence that captures its logical character—a view which he associates with Davidson. The logical character of a sentence is to be exhibited in the way it forms patterns with other sentences in virtue of which arguments are logically valid. The doctrine of logical form, as Szabó calls it, is the thesis that valid arguments can be sorted into the mutually exclusive categories of factually, lexically, and formally valid arguments, in accordance with whether their validity is explained by appeal to matters of fact (Alex is a father; Alex has a Y-chromosome), lexical meaning (Alex is a father; Alex has a child), or logical form (Alex is a father; Alex is a father or a mother). Szabó's strategy is to argue that the proponent of logical form cannot defend this distinction against two triviality objections. The first is that the validity of all inferences depends on facts. The second is that the validity of many formally valid inferences depends on definitions of logical words. Szabó considers a defense that appeals to forms

of semantic competence to distinguish between factually, lexically, and formally valid arguments. This is to give the distinction an epistemic basis. One may fail to recognize the factual validity of an argument while being fully competent in the language. If one understands the words in a lexically valid argument, then one will recognize its validity; if one recognizes the validity of an argument from formal features alone, then it is formally valid. Szabó construes competence in a language as a matter of tacit knowledge of a truth theory for it, and objects that this makes logical form relative to the deductive theory of the truth theory. In response to the maneuver we describe above that would seek to characterize sameness of logico-semantic form in terms of what all proofs in any interpretive truth theory have in common, Szabó protests that we arrive at a version not of logical form but of semantic form. This would be a problem, however, only if the notion of logical form at issue in Davidson is distinct from that of semantic form. It has been our view that it is not distinct, and that in this Davidson has simply been following a tradition that stretches back to Russell. (See Lepore and Ludwig (2002; 2007, ch. 7) for further discussion.)

“A truth predicate in the object language”

William Lycan

William Lycan proposes a novel way of responding to the problem of the semantic paradoxes for truth-theoretic semantics. Lycan’s goal is to describe how a natural language could work that would enable the application of a truth theory to it even though it contains its own univocal truth predicate. (It would be a further, and empirical question, whether natural languages do work in the way proposed.) The proposal Lycan makes is a variant of the idea of adopting an hierarchical approach to the language, but with a clever twist.

A standard hierarchy approach would introduce an ordered series of truth predicates for a language that can only take arguments of appropriate types (matching a predicate with a term denoting an argument inappropriate for its type would result in a category mistake, and the sentence would be said to be ungrammatical). “true₀” would apply to sentences which have no semantic vocabulary in them. Call these 0th-order sentences. Sentences in which “true₀” is predicated of a 0th-order sentence would be 1st-order sentences. “true₁” would apply to sentences at the 0th or 1st order, and sentences containing it would be 2nd-order sentences. In general, an *n*th-level truth predicate would apply to sentences of levels *n*–1 or lower. A problem for this approach is that it requires that we suppose that ordinary language truth predicates are, implausibly, infinitely ambiguous; and if the sense of each must be learned independently of the others, it would follow that this could not be the correct account of the semantic predicates of a natural language like English, because it would make it unlearnable by finite beings.

Lycan proposes that we treat the truth predicate as univocal, but as relativized to a language, and that we can think of English as in effect a series of languages. O is (as we may put it) the part of English which contains no semantical predicates. M₁ is the first

metalinguage, which embeds O but includes sentences containing "true" when its language argument is O. And so on. Lycan's trick is to restrict the grammatical sentences of English to those in which the implicit position for the language to which the truth predicate is relativized is restricted to the language immediately below the language in the hierarchy of which sentences with the truth predicate with its relativization is a component. Then when we consider the Liar sentence, "L is not true," we have to ask (a) what language it is in, and (b) to what language the truth predicate is relativized. L is not in O, because it contains "true." If it is in M_1 , then it is equivalent to "L is not true in O." Since L is not a sentence of O, this is not true. But this no more threatens paradox than does "This sentence is not true in German." Similarly, for any other legitimate interpretation of "is true" in L. Lycan shows how to extend this idea to quantified sentences such as "Everything John said last night is true" where we do not know what John said, and to both contingent and cyclical Liars.

"Swampman, response-dependence, and meaning"

Nathaniel Goldberg

Nathaniel Goldberg argues that the tension between Davidson's Swampman thought experiment (Davidson 2001k, p. 19) and his view that the standpoint of radical interpretation is the fundamental standpoint from which to investigate meaning and thought is an expression of two different response-dependent accounts of meaning.

In the Swampman thought experiment, Davidson asks us to imagine a human body which is a duplicate of his created when lightning strikes a tree near which he is standing. His body is destroyed. His duplicate, Swampman, is created out of different molecules, and now stands nearby. Davidson argues that the Swampman does not have any thoughts at all, and hence that a requirement for thought is that a subject have an appropriate history of interactions with his environment. The tension with taking radical interpretation to be the fundamental standpoint from which to investigate meaning and thought comes to this: from a complete description of the dispositions of Swampman to respond to things in the sorts of environments he can access, a radical interpreter (modulo the possibility of radical interpretation) could construct an interpretation theory for Swampman. If there is no more and no less to meaning than what is discoverable from the standpoint the radical interpreter, then we should conclude that Swampman does have a language and thoughts after all.

Goldberg argues first that Davidson has two response-dependent accounts of meaning. The concept of being funny is response-dependent in the sense that identifying what it subsumes depends conceptually on how a certain group responds or would respond to them under certain conditions (by being amused). Goldberg argues that Davidson is committed to the concept of meaning being similarly response-dependent. The first of the two accounts falls out of taking the radical interpreter's standpoint to be conceptually fundamental. If it is, then whether someone means something by a certain utterance, and what he means, is something that can be recovered in principle by the radical interpreter, in virtue of the nature of his subject matter. Thus, Goldberg argues,

the concept of meaning is response-dependent because we understand it as something that is present only if a suitable observer responds or would respond to it by subsuming it under an appropriate judgment about its meaning. (This seems to threaten to deflate the concept of response-dependence, for any discoverable property would seem to be guaranteed to be one such that an appropriately equipped inquirer with adequate evidence would respond to it by judging an object which has it to fall under the concept. The trouble lies in thinking of the response that characterizes the relevant category of things as a *judgment* that the thing is in the category, as opposed to a response that does not involve subsuming something under the category in question, like amusement, or a sensation.) The second of the two accounts falls out of Davidson's account of language learning, and what Goldberg sees as his commitment to the thesis that there is no meaning that is not learned. In language learning, Davidson emphasizes the importance, in basic cases, of triangulation between teacher, learner, and a common cause of their common responses to the environment. Goldberg's idea is that for the learner to come to mean something by a word, she must respond as the teacher does (in response to the teacher—forming the baseline of the triangle) to something in the environment, though here it is less clear what the response in terms of which the classification of an item as having a certain meaning is supposed to be.

Goldberg uses these two accounts to explain the tension that shows up in what Davidson says about the Swampman case between taking the stance of radical interpretation as basic and requiring a history of interaction with the environment. Goldberg argues that the requirement that emerges in the Swampman case is actually just the requirement that a speaker has learned what his words mean, which requires in some range of basic cases having triangulating with a teacher on a common cause of common responses. This is in conflict with what seems to fall out of reflection on radical interpretation, according to which to mean something is to be such that a radical interpreter could justifiably come to think you meant that.

Goldberg's conclusion is the same as the one we reached: namely, that Davidson should drop the diachronic requirement on meaning and thought content because it is incompatible with the view that the standpoint of radical interpretation is fundamental. Where he differs with us is in seeing the diachronic requirement as rooted an independent commitment of Davidson's to learning having to have taken place for meaning. Here we think our difference with Goldberg comes down to how we understand the role of talk of triangulation in Davidson's discussions of language learning. As we see it, this is a matter of emphasizing the importance of triangulation in communication by focusing on its role in language learning, but it is not supposed to suggest that for words to have the meanings they do, or any meanings, we have to find a direct or indirect route back to learning words in what Davidson describes as basic cases. If we learn a language in the usual way, triangulation is necessary, as it is necessary for all communication, and so, on Davidson's view, for meaning. But that we acquire language in the usual way is not, we think, something Davidson regarded as a necessary

condition on meaning something by one's words. One bit of evidence for this is that he allowed, in seminars and discussion, that the Swampman would come to mean things by its utterances eventually when it had interacted enough with things in its environment. No doubt some of this would involve interactions with other speakers that involve common responses to common causes in what we might as well call communication contexts, but there would be nothing here that would count as learning or teaching in any ordinary sense.

Part II

“Knowledge and error: a new approach to radical interpretation”

Olav Gjelsvik

Olav Gjelsvik makes the interesting suggestion that the objections that we raised in Lepore and Ludwig (2005, ch. 15) can be met by enriching our conception of the evidence available to the radical interpreter while remaining faithful to Davidson's view that the objective, third-person point of view is conceptually and methodologically fundamental in investigating meaning and thought. The main idea emerges out of considering what is required for an adequate account of error from the standpoint of the interpreter of another, both in belief and in action. Gjelsvik argues that an adequate account of error in belief requires us to take into account the subject's evidential sensitivities, and in particular to take into account what the subject is in a position to come to know on the basis of observation, where we think of this as involving how things look to an observer (cf. Katherin Glüer's essay in connection with this). We want to take into account not just whether the subject is in a position to observe a thing, but also whether it looks to him to be a thing of the kind it is, which is a separate matter. The radical interpreter, then, Gjelsvik suggests, will not just correlate conditions in the environment with hold-true attitudes, but will first look to statements about what the subject sees, hears, and so on, where these statements do not presuppose anything about what concepts the subject brings the objects he sees, hears, and so on, under. We proceed, then, from this to attributions of how things look if the speaker is to know what it is that he sees, hears, and so on, and we assume that by and large he does know what it is that he sees, hears, and so on. And interpretation proceeds from this point with this constraint. In the case of error in action, Gjelsvik argues that it comes down to failure to successfully execute intentions. The parallel to seeing and hearing in the case of action is the expression of a repertoire of basic actions as expressions of knowing how to do things but not by doing something else. Even attempts at basic actions can fail. But we assume that by and large they do not. We identify basic actions, and assume that the speaker knows he is performing them. In this way, data about basic actions constrain further theorizing.

Thus, if we understand Gjelsvik correctly, the enriched evidential base is to include facts about what an agent perceives (in various modalities) and what events involving the agent are basic actions. We are to assume that the agent by and large knows *what* he

perceives, and *what* basic actions he performs. This then forms a constraint on interpretation that guarantees mostly true environmentally directed beliefs without having to appeal independently to Charity.

There is a point on which we think we may differ with Gjelsvik in our interpretation of Davidson's project. As we see Davidson's project, it is to say how one could justify an interpretation theory on the basis ultimately of purely behavioural evidence. As a practical matter, he helps himself to hold true attitudes, but he describes this as an intermediate stage. For Gjelsvik, as we read him, the project is less ambitious. It is merely to start from data that do not presuppose knowledge of the detailed contents of a speaker's propositional attitudes. Seen from that perspective, helping oneself to knowledge of what movements of the agent's are basic actions, and so intentional under some description, and to facts about what the agents perceives, does not violate any constraint on the project, and it is liable, as Gjelsvik says, to help constrain interpretations significantly (though it remains to be seen whether it removes all objectionable underdetermination). However, if the ultimate evidence available to the radical interpreter is described purely physically, then statements about what a speaker sees and hears and what movements are basic actions likewise must be established on the basis of more primitive sorts of evidence.

If the ultimate evidence must be described in purely physical terms, then we cannot help ourselves to what the speaker sees and hears or what events involving his body are actions. These facts must be reconstructed from the primitive data. But since decisions about what the speaker sees and hears and what events involving his body are actions must be fit into an overall view of him as an agent and speaker, we are still faced with the prospect that there will be many equally good ways of doing it. We can help ourselves, as we do in practice, to the assumption that human speakers find the same things salient generally that we do—that is, see, hear, notice, and bring under similar concepts, the same things in our shared environment, adjusting for position. But this violates the constraint that we construct our account wholly from the third-person point of view, for we rely on knowledge of what we see, hear, notice, and so on, and then project it onto others. It gives us a workable theory, of course. But it does not give us reason to think that it is the only one justifiable compatibly with the constraints and evidence, and the challenge then remains to make good, in light of the possibility of finding different systematic causes of differential responses to the environment in speakers, to show that the constraints do not leave unacceptable underdetermination. And we should keep in mind that in the general case we cannot assume that a speaker does share saliences or even sensory capacities with us.

Does it matter for Davidson whether he retreats, as we would put, from the very austere picture of what evidence is ultimately available to the radical interpreter to something like what Gjelsvik has in mind? We think it does, because it is connected with the commitment to seeing the concepts of the theory of interpretation as theoretical concepts in the sense we have sketched. Gjelsvik accepts that his approach gives up on this thesis. But we think the thesis is important to Davidson because it is

what justifies thinking that what must be assumed from the radical interpreter's standpoint if he is to succeed is constitutive of its subject matter. We do not say that the project of showing how to get from information about what a speaker perceives and what events involving him are actions together with other information about his environment to an interpretation of him will not yield insights into the interpreter's subject matter. But we do not think it will secure, for example, that it is constitutive of subjects that they know things about their environment, or about the minds of others. To this extent, then, we think the response that Gjelsvik offers to Davidson involves a reduction of ambition.

“Perception and intermediaries”

Kathrin Glüer

Kathrin Glüer takes up Davidson's claim that only a belief can be a reason for holding a belief, reinterpreting this as the claim that for the proposition that p to count as a reason to hold that q , one must believe that p . This makes the proposition that p available as a premise for the subject in reasoning, and hence something that can be cited in support of other propositions, and so for coming to believe them. This rules out (R), however, which seems counterintuitive, since we seem to cite what we experience as reasons for what we believe.

(R) Perceptual experiences provide their subjects with reasons for belief.

Davidson rejected (R) on the grounds that experience does not have propositional content. Glüer suggests a way of retaining (R) in a basically Davidsonian framework, developing the proposal against the backdrop of the contrast between Davidson's view and McDowell's.

McDowell holds that perceptual experience has propositional content and that it can bear a rational relation of support to beliefs by way of its providing an entitlement to believe what it presents as so. Glüer argues that while the notion of entitlement may be important in epistemology, it would be a mistake to think that it exhausts the sorts of rational relations between experience and belief to which we appeal. Glüer grants that experience can have propositional content. The trouble is that entitlement is factive: if one is entitled to the belief that p on the basis of a perceptual representation of its being the case that p , then it is the case that p . But we still want to explain what reason someone had for a perceptual belief by citing her perceptual experience even when the perceptual experience turns out not to be veridical. Against this idea, McDowell urges that if we picture the relation as purely internal, then the inference from the experience that p to the belief that p is not defeasible, and so does not accommodate the element of risk that belief on the basis of perceptual experience entails. Glüer counters that if we construe the form of experiential content to be different from that of belief, we can obviate this objection and treat them as providing defeasible support for belief. More specifically, Glüer proposes that perceptual experiences are a kind of belief but with contents of a special form—though this is not meant to reject the special sensory nature

of experiences. This allows us to hold onto (R). To allow for the autonomy of non-sensory perceptual belief and perceptual experience, and to allow perceptual experience to defeasibly support perceptual belief, Glüer suggests that the propositional contents of perceptual experiences be conceived of as ascribing phenomenal properties to objects—properties such as appearing blue and round, and so on. Thus, the contents of perceptual experiences are different from those of the corresponding perceptual beliefs. This allows one to retain the perceptual representation, though one refrains from adopting the corresponding perceptual belief. And it allows for a genuine defeasible step from the perceptual representation to the perceptual belief. Finally, construing perceptual experiences as phenomenal beliefs about objects allows the application of charity to them in the same way as to other beliefs.

“On Davidson’s view of first-person authority”

Bruce Aune

Bruce Aune takes issue with some points in our interpretation of Davidson’s argument for first-person authority, and he offers a Sellarsian explanation of first-person knowledge to counter our pessimistic view that there is unlikely to be a philosophically illuminating explanation of first-person knowledge. We will touch briefly here on just two of the issues that Aune raises.

In “First Person Authority,” as noted above, Davidson aims to explain the “asymmetry between attributions of attitudes to our present selves and attributions of the same attitudes to other selves” (Davidson 2001j, p. 3) by appeal to an asymmetry in knowledge of meaning. If you and I both know that I hold true “Wagner died happy,” but there is a presumption that I know what I mean by that while there is no such presumption that you do, then there is a presumption that I know what I believe, but no presumption that you do. There appear to be two main arguments which Davidson offers for the assumption that there is a presumption that I know what I mean by my words but no presumption that you do. The first is that I can always state what my words mean (or the interpretive truth conditions for them) by disquotation—“Wagner died happy” is true iff Wagner died happy—but you cannot do the same because when you disquote you use your words on the right-hand side, and there is no guarantee that they mean the same as what my words mean. We objected that there is no obstacle to your uttering the same sentence that I do in my language, even if you do not understand it, as one might repeat a sentence in Latin (*audaces fortuna iuvat*) that one does not understand. Hence, being able to produce a T-sentence that is true and interpretive does not suffice to show that you know what the mentioned sentence means. Aune objects that just uttering the sentence is not to assert it or to say something, and so not to (in that sense) give truth conditions for it, and, in one sense, we agree; for to use a sentence to express a belief, one has to know what it means. But then production of truth conditions in this sense would presuppose knowledge of meaning, not explain it, so this way of construing the idea would not help Davidson. Aune’s own objection to Davidson’s argument is that the production of

an interpretive T-sentence for a sentence of one's language cannot explain to one what one means by it. For if one did not understand it, merely producing a T-sentence for it would not suffice. We agree with this criticism, and to our eye it looks similar to the point that we aimed to make.

Davidson's second argument relies on the fact that for interpretation to succeed, a speaker must apply his words correctly to things in the environment. Together with the claim that to be a speaker one must be interpretable, this yields the conclusion that a speaker must be disposed to apply his words correctly to things in his environment. If we take this to be sufficient for the speaker to know what his words mean, then we reach the conclusion that speakers know the meanings of their words. The asymmetry is supposed to arise from there being no argument that guarantees that an interpreter knows the meanings of the speaker's words. An important issue which Aune raises for this argument is whether it is compatible with Davidson's externalism about thought content. The challenge is to explain how a speaker could be authoritative about what his sentences mean and what his thoughts are if their interpretive truth conditions are fixed by what states of affairs in his environment regularly prompt them. For then what his thoughts are about would seem to be determined by facts with respect to which he does not have any special authority. So if he has to come to know what the contents of his thoughts are by attention to what determines those contents, it would seem he is no better placed than anyone else, even if as a matter of fact he uses his words correctly in application to things in his environment.

"Davidson, first-person authority, and the evidence for semantics"

Steven Gross

Steven Gross is concerned with what kind of challenge, if any, the fact of first-person authority about what one thinks and means presents to Davidson's commitment (as we have argued) to the claim that (E) the evidence available to the radical interpreter exhausts the relevant semantic facts and the allied assumption that (PT) the concepts of the theory of interpretation have their content exhausted by their "application in the domain of evidence in a way that results in the content of the theories' theoretical claims not transcending their predictions about facts in the domain of evidence" (Lepore and Ludwig 2005, p. 255). We objected to (E) and (PT) on two, as we saw it, connected grounds. First, the character of our first-person applications of the concepts of the theory is *prima facie* not compatible with (PT). Second, (E) requires that it make sense from the point of view of the interpreter of another that he can confirm a correct interpretation theory for a speaker, but this requirement is not met. Gross argues that the first objection fails, that the second does not rely on appeal to first-person authority, and that, finally, there are sources of evidence (the sort drawn on in psycholinguistics and cognitive neuroscience) that challenge (E) which are neither first-personal nor available to the radical interpreter.

We will not do justice to Gross's intricate discussion here, but we will try to say a little more about how we were thinking of the challenge that first-person authority

presents to Davidson, and the connection between the argument for underdetermination and first-person knowledge. Gross suggests that in our discussion we do not sufficiently distinguish between two claims:

- (i) the evidence available to a radical interpreter suffices for his recovering all the semantic facts, and
- (ii) ... for someone to ascribe with warrant an attitude or meaning to a speaker or the speaker's words, he must do so on the basis of such evidence. (This volume, p. 229)

The distinction is applied to an argument we give (Lepore and Ludwig 2005, p. 228) which focuses on the role of perceptual experience in acquiring knowledge of the world around us. As Gross construes the argument, it relies on the assumption that if a concept is theoretical it can be applied with warrant only on the basis of the evidence relevant to theory in which it has its home. Since we, in coming to learn things about the world around us, apply *inter alia* the concepts of belief and representation, but not on the basis of behavioral evidence about ourselves, they cannot be theoretical. Gross's objection is that all (E) requires is (i), but what we require in the relevant argument is (ii) in addition.

Gross's construal of the argument, however, does not quite conform to the way we were thinking of it, and for this we must accept the responsibility. There is, we think, a general challenge to Davidson to make sense of first-person knowledge and authority from the third-person perspective. But in the argument in question, what we were focusing on was how it is that we come to know the putatively content-exhausting evidence (evidence, for short) in the case of others. The point we aimed to make was that our understanding of how we come to know the evidence presupposes that knowledge of our own mental states and perceptual experiences is epistemically prior to knowledge of that evidence. But this is not compatible with seeing that evidence as epistemically prior (as it must be on Davidson's view) to the application of those concepts. The point here is not so much that we apply the relevant concepts to ourselves, but not on the basis of the relevant evidence (or even with warrant—for there are conceptions of warrant which would make (E) and warranted first person ascriptions consistent)—but that we conceive of the domain to which the concepts apply as epistemically prior to the domain which is to count the fundamental source of evidence for their application.

Gross considers an argument that we also present later in this book, and finds it likewise wanting. The argument occurs in the context of our discussion of Davidson's attempt to provide an explanation of first-person authority. We conclude that his argument that first-person authority can be explained by invoking what must be assumed by an interpreter is unsuccessful, and that in any case his target explanandum is too narrow. We grant, on other grounds, that the assumption that a speaker knows by and large what he means and what he thinks is unavoidable in interpretation. We deny, however, that this fact provides an explanation of first-person knowledge of thought and meaning. We argue that non-inferential knowledge of one's own

thoughts is required by our concept of a rational agent, and we follow Davidson in holding that we identify attitudes only in agents, and then only in patterns that are largely rational, so that rationality is constitutive of thought. On this view, then, considerable non-inferential knowledge of one's own thoughts is necessary to have any thoughts at all. Gross concentrates on the following passage toward the end of this discussion:

[The connection between rationality and self-knowledge] seems to show that [(a)] our *a priori* conception of a rational agent is one of a being who has non-inferential knowledge of its own psychological states, and, if it is a speaker, the meanings of its sentences, and [(b)] who therefore must regard its attempts to interpret others as attempts to discover facts which are not exhausted by what is recoverable from observations of behavior; [(c)] if there are two possible assignments of meanings and attitude contents to a speaker's sentences and attitudes, from the interpreter's standpoint, the possibility that one is right and the other wrong remains open, [(d)] because the interpreter must recognize the possibility of a perspective on those thoughts and meanings which is not dependent on recovering them from behavioral evidence. (Ibid., pp. 368–9)

Gross construes the argument in the following way: (b) follows from (a) *because* (d) follows from (a), (c) follows from (d), and (b) follows from (c) and the assumption that there is indeterminacy. Gross's objection is that the claim that (c) follows from (d) begs the question against Davidson. This is not quite how we intended this passage to be read. Rather, we intended to be saying: [a], therefore [b]. Furthermore, [c] because [d], where [d] itself, as Gross says, is supposed to rest on [a]. But in any case, if Gross accepts that [d] follows from [a], then it looks as if [PT] must be given up, whether or not [c] follows from [d], for [d] says explicitly that the interpreter must recognize a perspective on the thoughts of his subject—namely, the subject's own—that is not dependent on recovering them from behavioral evidence. What do we mean by saying that the speaker's (or anyone's) own perspective on his thoughts is not dependent on recovering them from behavioral evidence? We mean that the deliverances of that perspective is not conceived of as being constrained by what one could discover solely on the basis of purely behavioral evidence available to an arbitrary interpreter. But this is not compatible with [PT].

We suppose that, on this reading, Gross will deny that [d] follows from [a] and that [b] follows from [a] alone as well. We think this is where the issues lie, and that it merits more discussion than we gave it. But it is not possible to address it adequately in the space here, so we want to pass on briefly to the question of the relation of the first-person perspective to the argument for underdetermination. Gross argues that the argument we give against indeterminacy in favor of underdetermination, which would undermine (E), does not essentially draw on first-person authority. There is a sense in which we agree, for it need not advert to the asymmetry in epistemic authority between interpreter and speaker in any direct way. But there is a connection. It is important, we urged, that indeterminacy in interpretation make sense from the point of view of the radical interpreter himself. That point of view is, Davidson is committed

to saying, conceptually fundamental. The radical interpreter aims to develop an interpretive truth theory for the speaker. In doing so, he uses his own language in stating what the interpretive truth conditions of a speaker's sentences are. It is in virtue of his knowledge of what his sentences mean (or even of what their truth conditions are, whatever that comes to, as long as it involve knowing something about their semantic properties) that he is in position to see that there are different and *incompatible* theories that are compatible with all of the behavioral evidence. Gross is right in saying that it is enough that there be different incompatible interpretation theories—incompatible in the sense that they cannot all be (without qualification) correct, to show that we have underdetermination and not indeterminacy. But if this did not emerge from the point of view of the radical interpreter, we would not have shown, as we have hoped to, that the project is not coherent in its own terms.

“Davidsonian holism in recent philosophy of psychiatry”

Marga Reimer

Marga Reimer's contribution is concerned with a series of criticisms of Davidson based on reflection on psychiatric delusions. The critics hold that Davidson's views about attitudinal attributions and are incompatible with standard descriptions of cases of psychiatric delusions. Reimer argues that what she calls Davidson's holism—the view that attitudes can be intelligibly attributed only in largely rational patterns, when understood correctly—is in fact compatible with the relevant psychiatric phenomena. The central complaint against Davidson is that delusions do not always integrate with the “sort of rich pattern required, on the Davidsonian view, for the individuation and intelligible attribution of mental content” (this volume, p. 257).

Reimer argues, first, that there is no conflict between saying that thinkers are largely consistent and yet delusional. She says that a deluded thinker who, for example, thinks she can control the weather by mental command, does not simultaneously believe that the standard assumptions we make about the weather—that they are not under human control, that thunderstorms form suddenly in fronts and are subject to chaotic behavior—are universal generalizations: rather, the subject of such a delusion thinks these things hold generally but that she is an exception—in the same way that an otherwise scientific chemist who believes Christ turned water into wine at the marriage feast in Cana of Galilee may likewise make his beliefs consistent. This does of course entail that she (and the chemist) has a false belief. But that we have false beliefs is not something Davidson denies. He denies only that we can be massively in error about the world and in our general beliefs. And Reimer suggests that there is no reason to think that, by and large, subjects of psychiatric delusions do not have true beliefs. It might be pointed out that they generally do not have much trouble negotiating their environments, or talking to others about their delusions, and weaving a story about things they have seen into their narratives. On the question of whether those subject to serious psychiatric delusions integrate their delusions into a coherent pattern of beliefs, Reimer argues that there is no reason to think that they do not: for they act on their delusions,

both in what they seek (including information) and in what they avoid, and they tell elaborate and internally coherent stories justifying what they claim. Even in cases in which a thinker appears to have inconsistent beliefs, Reimer argues that there need be no conflict with Davidson's view. First, often on closer examination it is hard to find an outright synchronic inconsistency, and diachronic inconsistency is not *ipso facto* irrational, unless any time one changes one's mind one is irrational. Second, when someone actually says, for example, that she is both a man and a woman, or one person and simultaneously a distinct person, we feel pulled to question whether she could really believe that, which is what Davidson's overall view would predict. In short, Reimer sees little reason to locate any incompatibility between Davidson's holism about the attitudes and delusional thinking: the criticisms founder both on misinterpretations of the data provided by cases and on misinterpretations of Davidson's views.

“Taking back the excitement: construing ‘theoretical concepts’ so as to avoid the threat of underdetermination”

Richard N. Manning

Richard Manning takes issue with the argument we present for the impossibility of radical interpretation (Lepore and Ludwig 2005, ch. 15). As we have noted, the possibility of radical interpretation—that is, successful interpretation of any speaker from the standpoint of radical interpretation—is a central assumption in many of Davidson's philosophical projects. It is central to his project for illuminating meaning, and crucial to his rejection of traditional skepticism, to his claim that belief is by its nature veridical, to his argument that we can know the minds of others, and to his account of first-person authority. If radical interpretation is not possible, then Davidson's projects in these domains collapse. Manning aims to retain the excitement of Davidson's project by providing a construal of the notion of theoretical concept that can sustain the claim that the different theories at which the radical interpreter can arrive represent genuine indeterminacy.

We hold that Davidson takes the concepts of the theory of interpretation, including those of propositional attitude psychology, to be theoretical in the sense that their content is to be exhausted by the role they play in accounting for the relevant data—in this case the evidence available from the standpoint of the radical interpreter. We make two connected points against this. The first is that from the point of view of the radical interpreter, the evidence genuinely underdetermines interpretation theories of another speaker, in the sense that the radical interpreter recognizes that they attribute different interpretive truth conditions to the same sentences. The second is that this conception cannot accommodate the character of first-person attributions of these concepts.

Manning suggests that “a concept is theoretical if its content can be exhaustively characterized by appeal to the consequences of its playing its role in fulfilling the purposes for which the theory in question is constructed” (this volume, p. 275). This is to allow that they may have uses outside the context of the theory, but to restrict their content to their role in fulfilling the purposes of the theory. For a theory of

interpretation, Manning says that the purpose is “facilitating communication” in a “suitably deflated, practical” sense (*ibid.*, p. 276). The response to the argument against indeterminacy that we present relies on giving communication this “deflated, practical” sense in specifying the purpose of the theory. The idea is to reduce the ambition of an interpretation theory to enabling “the practical business of linguistic interaction to go smoothly” (*ibid.*, p. 281). The practical business of linguistic interaction, however, is not supposed to extend to the project of getting right what the speaker means by what he says. So an interpreter may have two theories that equally suffice for these purposes, though in fact they attribute different interpretive truth conditions to the same sentences.

This concession, however, appears to be a retreat from the claim that the concepts of the theory have their content exhausted by their roles in fulfilling the (deflated) purposes of the theory. For how can that claim be reconciled with the claim that different theories deploying these concepts in fact ascribe different, and incompatible, properties to sentences of the speaker’s language? At this point, Manning suggests that we have misconstrued Davidson’s assumption that radical interpretation is possible as a premise in his overall argument. He suggests that instead it is “a *recommendation* on Davidson’s part for how we ought to think about linguistic concepts and the attitudes” (*ibid.*, p. 285); if we have the idea right, the suggestion is that we should reinterpret our ordinary discourse about meaning and propositional attitudes so that the concepts they express in fact have their contents exhausted by their role in facilitating communication in a suitably deflationary sense. Putting questions of interpretation aside, however, this leaves us with the question of whether the excitement of Davidson’s philosophical project can really be rescued in this way. For when we consider what Davidson has to say about traditional philosophical problems on this interpretation, we have to interpret the words differently than in the statement of the problems, which threatens to make their engagement with those problems merely verbal.

References

- Carnap, Rudolf. 1947. *Meaning and Necessity*. Chicago: University of Chicago Press.
- Davidson, Donald. 1975. Thought and Talk. In S. Guttenplan (ed.) *Mind and Language*. Oxford: Oxford University Press, pp. 7–23.
- 1985. A New Basis for Decision Theory. *Theory and Decision* 18: 87–98.
- 1990. The Structure and Content of Truth. *The Journal of Philosophy* 87 (6): 279–328.
- 2001a. Theories of Meaning and Learnable Languages. In *Inquiries into Truth and Interpretation*. New York: Clarendon Press, pp. 3–16. Original edition, 1966.
- 2001b. The Logical Form of Action Sentences. In *Essays on Actions and Events*. New York: Clarendon Press, pp. 105–21. Original edition, 1967.
- 2001c. Truth and Meaning. In *Inquiries into Truth and Interpretation*. New York: Clarendon Press, pp. 17–42. Original edition, 1967.
- 2001d. On Saying That. In *Inquiries into Truth and Interpretation*. New York: Clarendon Press, pp. 93–108. Original edition, 1968.

- 2001e. Radical Interpretation. In *Inquiries into Truth and Interpretation*. New York: Clarendon Press, pp. 125–40. Original edition, 1973.
- 2001f. Thought and Talk. In *Inquiries into Truth and Interpretation*. New York: Clarendon Press, pp. 155–70. Original edition, 1975.
- 2001g. Reply to Foster. In *Inquiries into Truth and Interpretation*. New York: Clarendon Press, pp. 171–9. Original edition, 1976.
- 2001h. Empirical Content. In *Subjective, Intersubjective, Objective*. New York: Clarendon Press, pp. 159–76. Original edition, 1982.
- 2001i. A Coherence Theory of Truth and Knowledge. In *Subjective, Intersubjective, Objective*. New York: Clarendon Press, pp. 137–53. Original edition, 1983.
- 2001j. First Person Authority. In *Subjective, Intersubjective, Objective*. New York: Clarendon Press, pp. 3–14. Original edition, 1984.
- 2001k. Knowing One's Own Mind. In *Subjective, Intersubjective, Objective*. New York: Clarendon Press, pp. 15–38. Original edition, 1987.
- 2001l. Three Varieties of Knowledge. In *Subjective, Intersubjective, Objective*. New York: Clarendon Press, pp. 205–19. Original edition, 1988.
- 2004. *Problems of Rationality*. Oxford: Oxford University Press.
- 2005a. A Nice Derangement of Epitaphs. In *Truth, Language, and History*. Oxford: Oxford University Press, pp. 89–108.
- 2005b. *Truth, Language and History*. Oxford: Clarendon Press.
- Dummett, Michael. 1975. What is a Theory of Meaning? In S. Guttenplan (ed.) *Mind and Language*. Oxford: Oxford University Press, pp. 97–138.
- 1991. *The Logical Basis of Metaphysics*. Cambridge: Harvard University Press.
- Frege, Gottlob. 1997. On Sinn and Bedeutung. In *The Frege Reader*, ed. M. Beaney. Oxford: Blackwell Publishers, pp. 151–71. Original edition, 1892.
- Lepore, Ernest, and Ludwig, Kirk. 2002. What Is Logical Form? In *Logical Form and Language*, ed. G. Preyer. Oxford: Clarendon Press, pp. 54–90.
- 2005. *Donald Davidson: Meaning, Truth, Language, and Reality*. Oxford: Oxford University Press.
- 2007. *Donald Davidson: Truth-theoretic Semantics*. New York: Oxford University Press.
- 2011. Truth and Meaning Redux. *Philosophical Studies* 154(2): 251–77.
- Soames, S. 2008. Truth and Meaning: In Perspective. *Truth and Its Deformities: Midwest Studies in Philosophy* 32: 1–19.

