

The Relevance of Radical Interpretation*

Jonathan Ellis

There is an interesting parallel to be drawn between the method of Davidson's philosophy and what Davidson's philosophy is about, or purports to show. Throughout his writings, considerations of language and interpretation are what consistently and ultimately yield his varied and bold views on the mental. Perhaps the most central of those views—and here is the parallel—is that mental states, or thought in general, emerge only in the context of communication and interpretation. The mind is born of interpretation, or language, as Davidson's own philosophy of mind is born of his philosophy of language.

Drawing the parallel reveals what I think is a persistent bias in Davidson's work. Asking for the source of that parallel can lead one on a journey through Davidson's assorted writings in pursuit of the origin, and thus foundation, of those views of the mental which are to spring so directly from his remarks on interpretation. What that inquiry reveals, what the parallel calls attention to is a consistent confidence on Davidson's part in the primacy and possibility of interpretation of the mental. An appreciation of the dependence of Davidson's arguments upon a prior view about interpretation is important. Many philosophers would not endorse such a view, which is perhaps one reason that exploring interpretation continues to be an uncommon approach for investigations into the nature of the mind. The very relevance of interpretation for understanding the nature of the mind—or at least the deep relevance Davidson sees in it—turns on a prior and controversial view.

* Forthcoming in *Engaging Davidson: New Perspectives on the Philosophy of Donald Davidson* (ed. J. Malpas), MIT Press, Cambridge, Mass. Other contributors include Lee Braver, Gordon G. Brittan, Jr, Sharyn Clough, Robert Dostal, Jonathan Ellis, Dagfinn Føllesdal, Barbara Fultner, Dieter Henrich, David Couzens Hoy & Christoph Durt, Jeff Malpas, Richard N. Manning, Giancarlo Marchetti, Mark Okrent, Gerhard Preyer, Louise Roska-Hardy, Bjørn Ramberg, Richard Rorty, Frederick Stoutland, Stephen Turner, David Vessey, Samuel C. Wheeler III, and Yang Xiao.

What is that view, and is it supported by Davidson's work? That is the focus of this paper. When we look carefully at Davidson's body of work, explicit justification or explanation for the view we will identify is difficult to find. Moreover, the little justification to be found appeals itself to interpretation, an appeal that we will see depends on the very view about interpretation in question. Much of Davidson's philosophy depends upon—starts with—a view of the mind that is unsupported by anything Davidson says about interpretation, or about anything else. Our investigation will not suggest that this view is incorrect (nor will it suggest that it is correct). But it will raise interesting questions about the force of Davidson's arguments and, more generally, about the starting point in Davidsonian philosophy.

1. Davidson's Theory of the Belief

Many of the substantial conclusions Davidson draws about the nature of mental phenomena (for him, the propositional attitudes) emerge largely through discussions of belief in particular. And invariably, his method is to "adopt the stance of a radical interpreter when asking about the nature of belief."¹ Two such conclusions are his anomalous monism and his claim that thought is possible only for creatures in communication with others. What leads Davidson to proclaim his principle of the anomalism of the mental and, in turn, to establish anomalous monism is the purported fact that the correct attribution of beliefs requires sensitivity to normative principles of coherency and rationality, which he infers from the constraints upon a successful radical interpreter.² His conclusion that thought is possible only for creatures in communication with others stems from other, related constraints upon the radical interpreter. If

¹ D. Davidson, "A Coherence Theory of Truth and Knowledge", in E. LePore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (Oxford: Basil Blackwell, 1986), p.315. In the sentence I quote, Davidson is describing the method of the particular argument he is about to provide for his claim that belief is "in its nature veridical."

² D. Davidson, "Mental Events", in his *Essays on Actions and Events* (Oxford: Oxford University Press, 1980).

a radical interpreter is to be successful, he argues, she must take her subject's beliefs often to be about those objects and events that she thinks cause those beliefs. One's beliefs, Davidson ultimately concludes, are in the most basic cases about precisely those objects and events in one's environment that cause those beliefs. For any belief, however, there are many events that play a causal role in its coming to be. What determines which of those many causes is the content of the belief, according to Davidson, is a process of triangulation involving interpreter(s) and subject.³

While there is no one article in which he details all aspects of his conception of belief (a non-representational and, as he calls it, "antisubjectivist"⁴ view), a detailed composite can be put together from the wide variety of his articles that concern belief in one way or another.⁵ According to Davidson, we are no longer to think of beliefs either as entities or as involving objects—propositions, representations, sense-data—that the mind somehow contemplates or "grasps."⁶ That is an old and pernicious picture of the mind, what he calls "the myth of the subjective,"⁷ which we are to abandon at once. Rather, we are advised to think of beliefs as "constructs"⁸ of our own theories of interpretation and action. Beliefs have "jobs"⁹ to perform in the interpretation and understanding of the behavior of others (and ourselves) and are thus

³ See D. Davidson, "Three Varieties of Knowledge", in A. Phillips Griffiths (ed.), *A.J. Ayer: Memorial Essays. Royal Institute of Philosophy Supplement* (Cambridge: Cambridge University Press, 1991), pp.153-166, and D. Davidson, "Epistemology Externalized", *Dialectica*, 45 (1991), pp.191-202.

⁴ D. Davidson, "The Myth of the Subjective", in M. Krausz (ed.), *Relativism: Interpretation and Confrontation* (Notre Dame: University of Notre Dame Press, 1989).

⁵ Davidson's papers involving issues about belief span more than two decades. It is perhaps best to regard his many ideas over those years as constituting an evolving conception of belief, and not a position that was entirely there from the start.

⁶ Davidson, "The Myth of the Subjective", p.171, and D. Davidson, "Knowing One's Own Mind", *Proceedings and Addresses of the American Philosophical Association* 60 (1987): pp.453-455.

⁷ Davidson, "Knowing One's Own Mind", p.456.

⁸ D. Davidson, "Belief and the Basis of Meaning" (1974). Reprinted in *Inquiries into Truth and Interpretation* (Oxford: Oxford University Press, 1984), p.146.

⁹ *Ibid.*, p.147.

¹⁰ *Ibid.*, p.153.

“built”¹⁰ by us to fulfill such roles. Beliefs, as he says, are then “best understood in their role of rationalizing choices or preferences.”¹¹

Since an important condition of successful interpretation is that the subject be rendered rational (according to the interpreter), an essential feature of belief is that its attribution always be held to normative constraints of rationality and consistency with respect to the totality of one's actions and utterances. For Davidson, any set of beliefs at all must largely abide by such normative principles. To the extent that there exists more than one theory about a person's set of beliefs that optimize his rationality, i.e., to the extent that a number of theories explain his behavior equally well, it is an indeterminate matter as to which theory is correct about his beliefs. It is not that the evidence is insufficient to tell which theory is right—it is not epistemological in that sense—but rather that each of those theories, while attributing different beliefs to the person, simply captures everything of relevance there is to capture.¹² The propositions we might use to express the objects of particular beliefs, Davidson suggests, might thus be “overdesigned”¹³ for their jobs. That beliefs are always as unique and fine as our ways of describing and attributing them might suggest is a false thought, and one only further encouraged by the view that beliefs have propositions as their objects.¹⁴

According to Davidson, then, beliefs are primarily to be construed as elements or “constructs” of a third-person explanation of behavior. That is not to say that we do not really have beliefs; we do. Having been built in the context of interpretation, the created predicates, Davidson believes, are then objectively true or false of us. What have been constructed are predicates that capture aspects of the complicated structure of one's behavior and dispositions to behavior. That is why, like meaning, belief is public, open to view and, in principle, entirely

¹¹ Ibid., p.147.

¹² Ibid., p.154.

¹³ Ibid., p.147.

interpretable. Some philosophers thus attribute to Davidson “constitutive interpretationism,” according to which a statement of what it is for *S* to believe that *p* makes essential reference to the idea of *S*’s being interpretable as believing that *p*.¹⁵ On another formulation of constitutive interpretationism, *S*’s believing that *p* “consists in” or is “nothing more than” *S*’s being such that *S* is *interpretable* as believing that *p*. Constitutive interpretationism is stronger than the claim, also attributed to Davidson, that *S* believes that *p* if and only if *S* is interpretable as believing that *p*.

The public nature of belief can also be seen in part from the fact that what in the most basic cases determines the contents of our beliefs are, for Davidson, the objects and events that actually cause those beliefs.¹⁶ Those are what the beliefs they effect are *about*. An interpreter who had complete knowledge of all of a subject’s potential behavior and the circumstances under which it occurred would be in the position to know everything a speaker believes.¹⁷ Belief is thus no longer “subjective” or private in the way it is sometimes thought to be; but rather “as a private attitude it is not intelligible except as an adjustment to the public norm provided by language.”¹⁸ Not only is belief public in this epistemological sense concerning the fully informed interpreter, but there is the even further conclusion, mentioned above, that belief in general is only possible for creatures already engaged in communication; thought emerges only in the context of interpretation, and it is thus essentially social.

¹⁴ Ibid.

¹⁵ See W. Child, *Causality, Interpretation, and the Mind* (Oxford: Oxford University Press, 1994), pp.48-53.

¹⁶ Davidson, "The Myth of the Subjective", p.171.

¹⁷ Davidson, "A Coherence Theory of Truth and Knowledge", p.315.

¹⁸ D. Davidson, "Thought and Talk", in his *Inquiries into Truth and Interpretation* (Oxford: Oxford University Press, 1984), p.170.

Such a conception of belief stands in contrast to more traditional or “Cartesian” conceptions of the mind, according to which beliefs are “internal” entities or involve mental objects or representations the mind grasps. For those who view belief in this way, whether a subject has a given belief will seem quite a determinate matter of fact. And while it may turn out that often the set of one's beliefs does satisfy normative ideals of rationality and consistency, that it does is not always, on such accounts, an essential mark of the mental. Nor on such conceptions is the mental typically thought to be essentially interpretable, or entirely “public.” There is rarely reason on such models to suppose that one's having a belief presupposes one's being in communication with others. Far from being essentially social, belief often seems essentially individual. And far from being essentially public, it often seems essentially private.

It is worth emphasizing that Davidson's picture of belief is not only what grounds his influential claims in the philosophy of mind (e.g., concerning anomalous monism and the social character of thought), it is what fuels many of his most provocative and ambitious claims in epistemology. We are to avoid skepticism of the senses, we are told, once we realize that the contents of beliefs are in general what cause those beliefs.¹⁹ Knowledge of other minds is also easily secured given that thought is essentially social. Abandoning the idea of objects of the mind is helpful for a number of reasons. Besides ridding us of a number of traditional worries concerning representation (such as how such objects represent, what it *is* they represent, and how the mind relates to those representations), discarding that view is what Davidson believes allows an externalist such as himself to maintain first person authority.²⁰ Additionally, he claims

¹⁹ Davidson, "A Coherence Theory of Truth and Knowledge".

²⁰ Davidson, "Knowing One's Own Mind".

that that picture of the mind is the source of the “deep mistake” we have long made in distinguishing scheme and content.²¹

Davidson's conception of belief is thus one reason his philosophy has such breadth. However, I am interested not in these larger conclusions but in Davidson's conception of belief itself, and in particular, in his appeal to radical interpretation in support of this conception. It is not clear precisely how this appeal is to work. The formidable consequences Davidson hopes to attain from his conception of belief only make such an inquiry all the more worthwhile.

2. Radical Interpretation

As I have said, Davidson portrays many aspects of his conception of belief as ultimately stemming from considerations about radical interpretation. It will be helpful, then, to begin with a description of the situation of radical interpretation and some of the initial conclusions Davidson draws from it. She who finds herself in the situation of radical interpretation is faced with the task of forming a theory of interpretation for a given speaker (or speakers)—a theory of what that speaker means by his utterances—without knowing anything at all to begin with about what that speaker means. Neither, though, is the interpreter able to make use of evidence involving the speaker's beliefs or intentions, for, as Davidson is apt to point out, she could not have evidence of these without already having a working theory of her speaker's meanings. This is because of the crucial connection between meaning and belief: to be justified in attributing certain meanings to a speaker we must know something about what he believes and intends, and to be justified in attributing certain beliefs and intentions to him, we must know something about what he means by his words. This insight we see put to use in a number of ways in Davidson's philosophy. Not only has the task of the interpreter expanded—now she must

²¹ Davidson, "The Myth of the Subjective", p.171.

simultaneously find a theory of belief (and desire and intention) as well as a theory of interpretation—but her evidential base has become more stark. All she has to go on are the speaker's actual utterances and the circumstances under which they are uttered.

Davidson takes it for granted that in addition to her speaker's actual utterances and the circumstances under which they are uttered, an interpreter would also be able to detect when her speaker holds particular utterances or sentences to be true. The difficulty, however, is that one holds a sentence to be true both because of what one means by it and because of what one believes. Davidson calls the attitude of holding true a “vector” of meaning and belief, neither of which our interpreter knows at this point. Suppose, for example, that the interpreter detects that her speaker holds the utterance “Gavagai” to be true in all and only those circumstances in which the interpreter sees a rabbit passing in front of the speaker. Without reason to think that in these circumstances the speaker *himself* believes that there is a rabbit in the vicinity, there is little reason to think that “Gavagai” for him has anything at all to do with rabbits. But at this juncture a radical interpreter does not know anything about her speaker’s beliefs. Disentangling the effects of meaning and belief becomes a central task of the interpreter.

From Frank Ramsey's method in decision theory, Davidson adopts the idea that in order to disentangle these effects, our interpreter must find a legitimate way to hold one of the two factors fixed while solving for the other.²² That way is at hand, Davidson believes, once we see that an interpreter is justified in assuming that most of some central core of her speaker's beliefs are in agreement with her own. What justifies such an assumption is the fact that disagreement, and the identification of belief in general, are possible only against a background of massive agreement. Appealing to Quine, Davidson writes, “Quine's key idea is that the correct

²² F. P. Ramsey, "Truth and Probability," in his *Foundations of Mathematics* (New York: Humanities Press, 1950).

²³ D. Davidson, "The Structure and Content of Truth", *Journal of Philosophy* 87 (1990), p.319.

interpretation of an agent by another cannot intelligibly admit certain kinds and degrees of difference between interpreter and interpreted with respect to belief.”²³ Such agreement encompasses not only the beliefs themselves but the principles of rationality and logic that connect and ground those beliefs. Without assuming that her speaker's beliefs adhere to a similar rational structure as her own, there would be nothing to guide her construction of her theory. The strategy of interpretation, then, is to assume, until proven wrong, that the speaker believes (and reasons) as the interpreter does. Although that assumption might be wrong, assuming agreement is at least always a good strategy. For if such agreement is not generally there, interpretation was doomed from the start, according to Davidson. Such agreement is thus both a finding of, and a requirement for, successful interpretation.

The interpreter can now justifiably begin to correlate sentences that her speaker holds true on certain occasions with that which is going on in the environment at those times. By looking at what in the environment the interpreter believes is systematically causing her speaker to hold such sentences true, the interpreter can begin to determine what her subject means by his words. For the interpreter can now assume that what she herself believes is going on in the environment at those times is just what her speaker believes is going on. Successful interpreters, according to Davidson, must thus generally take their subject's most basic utterances and beliefs to be about just those things in the world that cause those utterances and beliefs. Once the interpreter has a handle on those of her speaker's utterances that concern the more observable goings-on in the environment, she can then proceed to attempt to determine, by way of that which she has already learned, the meanings of those utterances that are not as directly tied to the observable environment.

The central role of belief, though, has not yet emerged. For belief has little role when speaker and interpreter are in agreement; the assumption of agreement is helpful precisely because it takes belief *out* of the picture. The interpreter's assumption of agreement, however,

will at some point render the speaker's behavior quite irrational, as there is bound to be considerable disagreement between them. For Davidson, the central role of belief is to render this behavior rational. And to do that, the interpreter attributes error. As Davidson says, "Error is what gives belief its point." And later:

Since the attitude of holding true is the same, whether the sentence is true or not, it corresponds directly to belief. The concept of belief thus stands ready to take up the slack between objective truth and the held true, and we come to understand it just in this connection.²⁴

It is helpful here to understand "the objective truth" as referring to what the interpreter *considers* to be the objective truth, i.e., as referring to what the interpreter believes. Belief comes in when the interpreter cannot preserve full rationality simply by adjusting her speaker's meanings.

That is the method of the radical interpreter, and it is from considerations about this method that Davidson often quickly draws conclusions about the nature of belief. He says, for instance (directly after outlining the procedures of the interpreter and certain ideas from decision theory), "Broadly stated, my theme is that we should think of meanings and beliefs as interrelated *constructs* of a single theory just as we already view subjective values and probabilities as interrelated constructs of decision theory."²⁵ Once we are to think of belief in this way, it becomes clear why it is sometimes an indeterminate matter as to whether one has a particular belief or not, and also why we are not to think of beliefs either as entities or as involving objects—propositions, representations, sense-data—that the mind somehow contemplates or grasps. Furthermore, not only are we to think of beliefs as theoretical

²⁴ Davidson, "Thought and Talk" p.170.

²⁵ Davidson, "Belief and the Basis of Meaning", p.146, italics mine.

“constructs,” but the construction of such theories is the very context in which we get our idea of belief in the first place. We “come to understand it” in precisely that context. “What makes a social theory of interpretation possible is that we can construct a plurality of private belief structures: belief is built to take up the slack between sentences held true by individuals and sentences true (or false) by public standards.”²⁶

3. The Relevance of Conditions for Successful Interpretation

As I have noted Davidson’s method is invariably to “adopt the stance of a radical interpreter when asking about the nature of belief.”²⁷ Why does Davidson adopt this method? Why does Davidson believe that investigating what one would do, even must do, in the unique and seemingly hypothetical situation of radical interpretation will shed light upon the nature of belief in general? When we start to inquire, explicit explanation on Davidson's part is difficult to find.

In one of Davidson's early articles on the topic, “Radical Interpretation,” where he says the most about its actual role in our lives, he implies that we find ourselves in this situation *every* time we understand what another person says. Radical interpretation is not applicable only to interpreters of speakers of a language entirely foreign to them. But rather, *all* instances of understanding what another speaker says are supposed to be instances of, or require, radical interpretation. He says,

The problem of interpretation is domestic as well as foreign: it surfaces for speakers of the same language in the form of the question, how can it be determined that the language is the same? Speakers of the same language can go on the assumption that

²⁶ Ibid., p.153. Which of these claims about belief Davidson intends to *infer* from his considerations about radical interpretation, and which he intends simply to make along with them is in some cases unclear. It often seems the former. Our investigation will be revealing in either case.

for them the same expressions are to be interpreted in the same way, but this does not indicate what justifies the assumption. *All understanding of the speech of another involves radical interpretation.*²⁸

However, Davidson later abandons the idea that all understanding requires radical interpretation. He writes more recently, “The approach to the problems of meaning, belief, and desire which I have outlined is not, I am sure it is clear, meant to throw any direct light on how in real life we come to understand each other. . . .”²⁹ The situation of radical interpretation is thus better thought of as the subject of a thought experiment.³⁰ The question is: If radical interpretation is not a situation we typically or ever find ourselves in, why should consideration of what we would do in that situation yield insights into the nature of belief?

In one of the few places where Davidson explicitly addresses the relevance of an appeal to radical interpretation, he writes (after that which I just quoted): “I have been engaged in a conceptual exercise aimed at revealing the dependencies among our basic propositional attitudes at a level fundamental enough to avoid the assumption that we can come to grasp them—or intelligibly attribute them to others—one at a time.” Investigation of the situation of radical interpretation, even if it is not meant to be an investigation of anything that actually happens, is intended to shed light upon the “dependencies” among the propositional attitudes that we do

²⁷ Davidson, "A Coherence Theory of Truth and Knowledge", p.315.

²⁸ D. Davidson, "Radical Interpretation", in his *Inquiries into Truth and Interpretation* (Oxford: Oxford University Press, 1984), p.125, italics mine.

²⁹ Davidson, "The Structure and Content of Truth", pp.324-325.

³⁰ Davidson's remarks here are in tension perhaps with his description elsewhere of the advantage of his arguments for externalism about content over those of Tyler Burge and Hilary Putnam. His own arguments, he says there, “do not rest on intuitions concerning what we would say if certain counterfactuals were true. No science fiction or thought experiments are required.” (Davidson, "Knowing One's Own Mind", p.450.

have, such as our particular beliefs and meanings. What we realize when we consider how one must go about radically interpreting a given speaker is how dependent one's correct attributions of certain of one's speaker's propositional attitudes are upon others of one's correct attributions of one's speaker's propositional attitudes. Often we do not realize these dependencies when we attribute beliefs on a daily basis, since in those cases we already know so many of the propositional attitudes of our speaker. We are hardly aware of our knowledge of those dependencies, and more importantly, of the role that our knowledge of them and that our sensitivity to principles of rationality play in our understanding of others. That is why it is sometimes easy to suppose that “we can come to grasp them—or intelligibly attribute them to others—one at a time.” Consideration of the radical interpreter helps us to appreciate and acknowledge these dependencies, principles, and relations which must exist and by which we must abide.

These dependencies have center stage in the prescribed method of the radical interpreter. The reason the radical interpreter must implement such a complex method in order to discern one's meanings and beliefs is that what the interpreter observes—the holding true of sentences under certain circumstances—is the product of at least two forces: meaning and belief. The crucial point is that, as we saw, we cannot ascertain what a speaker means by a given sentence that we know he holds true without already knowing the belief that the speaker has about the world that is the basis for his assent to the sentence. No matter what circumstances *we* believe obtain when he assents to that sentence, we cannot tell what he means by that sentence unless we know what circumstances *he* believes obtain. Likewise, we cannot ascertain from his assent to a particular sentence the belief that is the basis of his assent without knowing what he means by that sentence.

Untangling the effects due to meaning and belief is the goal of the method Davidson prescribes; it is what the method is invented to solve. The method is first to assume that our

speaker's beliefs are true. Holding belief fixed in that way, we can then solve for meaning. Subsequent adjusting of the theory to comport with the individual's particular idiosyncrasies only further brings out the normative principles and relations of rationality and consistency with which we inevitably endow our subject's attitudes.

However, the interdependencies of these propositional attitudes are emphasized only in consideration of what we must do to *ascertain* someone's propositional attitudes. The context in which these propositional attitudes are concluded to be interdependent is always an *epistemological* one. These aspects of the mental might not be brought out were we to consider what must be the case for someone to *have* a given belief, as opposed to considering what must be the case for one to *attribute correctly to someone* a given belief. From the fact that our correct attribution of someone's beliefs depends upon our correct attribution of what he means, it does not follow that his very having of those beliefs itself depends upon his very having of those meanings. For these considerations about correct attribution to yield such conclusions something more would be needed than that which has been said thus far.

Had a thesis such as constitutive interpretationism already been secured, belief's interdependency with meaning might follow. If a correct statement of what it is for *S* to believe that *p* makes essential reference to the idea of *S*'s being interpretable as believing that *p*, and if successful belief attribution always depends upon successful meaning attribution, and vice versa, then perhaps it would be harmless to consider belief and meaning to be interdependent. Reasoning in this way would require that we have reason to embrace this constitutive conception of belief, yet it is the source of that and other aspects of Davidson's conception of belief that is the subject of our inquiry.

Likewise, it does not follow solely from the fact that a successful interpreter must assume and attribute normative principles of rationality and consistency to her subject that belief is essentially normative in character, i.e., that it is in the nature of belief that one's set of beliefs

(and other attitudes) is always founded upon, and held together by, such normative principles of rationality and consistency. There may be good reason on *other* grounds, independent of considerations of interpretation, to suppose belief is essentially normative in character. Indeed, Davidson sometimes offers such grounds. My point here is only that such a claim about the nature of belief does not follow from consideration of interpretation alone. These considerations of interpretation so far entail nothing at all about what beliefs *are* or *are like in general*. They entail merely something about their correct attribution (as well as perhaps something about what they must be like *if* they are to be correctly attributed).

This point can be brought out in a different way. Those who disagree with Davidson's conception of belief could agree with everything Davidson says about the difficulties a radical interpreter would face, the procedures she would have to implement, and the constraints by which she would have to abide in order successfully to interpret her speaker. For the sake of example, consider the antiquated view—what Quine calls the “myth of a museum”—according to which meanings (and, in our case, beliefs) are inner items that glide across one's internal stage. This is perhaps close to what Locke or Hume thought, and it is an extreme version of the traditional “Cartesian” picture of the mind that Davidson wants to abandon. But even this position is not incompatible with Davidson's conclusions about the methods and constraints upon a radical interpreter. It is instructive to investigate what implications Davidson's considerations about radical interpretation have for such a view.

One who held such a conception of belief (call it the “museum conception”) might conclude about the plight of the radical interpreter much the same as Davidson concludes. For instance, he would most likely agree that a radical interpreter could not ascertain what a given speaker believes without assuming something about what that speaker means, and vice versa. For according to our museum theorist, one's meanings and beliefs consist in the appearance of certain mental entities before one's mind, and the interpreter would not have access to these

entities. The museum theorist would perhaps also thus agree that some procedure of holding one of the two fixed (e.g., by assuming largely shared belief) would have to be implemented; otherwise, an interpreter could not untangle the effect of the two. He might thus agree that successful interpretation always guarantees, and thus requires, massive agreement. The museum theorist could—and perhaps would—agree with all this, and yet still hold that what a belief *is* is the appearance in one's mind of an inner, mental entity.

The museum theorist would also likely agree that in order successfully to interpret her speaker, an interpreter would have to proceed upon the assumption that her speaker's attitudes adhered to normative ideals of rationality and consistency. Otherwise, there would be nothing to guide the interpreter's construction of her theory. Successful interpretation, he would acknowledge, must always render one's beliefs and meanings as largely satisfying normative ideals. The museum theorist might even hold that one's set of beliefs actually *does*, or even must, adhere to such normative principles. That claim is also not incompatible with a museum conception of belief, though perhaps it is not entailed by it.

Our museum theorist could therefore agree with much of what Davidson argues about the epistemological limitations and constraints upon the radical interpreter and about the procedures one would have to implement to radically interpret a speaker, yet not agree with, say, the constitutive conception of belief, or with the idea that beliefs are not inner items that glide across an internal stage, or the idea that it is sometimes a wholly indeterminate matter what beliefs one has, and so on. Concluding that an interpreter must construct a theory of belief in such a way does not commit one to very much at all about the nature of belief. What the constructed theory could do, for someone such as our museum theorist, is attribute determinate entities or objects that do appear in the speaker's mind.

The museum conception of belief is a confused one. But that only illuminates the explanatory status of radical interpretation all the more.

4. The Possibility of Interpretation

I have claimed that radical interpretation as considered so far entails nothing at all about what beliefs are, or are in general like, just something about their correct attribution, as well as something about what they must be like in the case in which they are to be successfully interpreted. Perhaps radical interpretation would show more about the general nature of belief were there reason to think that a creature's set of beliefs is always of a nature such that successful interpretation of it is possible. Davidson does emphasize the importance of the very possibility of interpretation. This passage from "The Structure and Content of Truth" from which I have already quoted addresses the issue more directly than any other. The bulk of the passage reads as follows:

I have been engaged in a conceptual exercise aimed at revealing the dependencies among our basic propositional attitudes. . . . Performing the exercise has required showing how it is in principle possible to arrive at all of them at once. Showing this amounts to presenting an informal proof that we have endowed thought, desire, and speech with a structure that makes interpretation possible. Of course, we knew it was possible in advance. The philosophical question was, *what* makes it possible?

What makes the task practicable at all is the structure the normative character of thought, desire, speech, and action imposes on correct attributions of attitudes to others, and hence on interpretations of their speech and explanations of their actions.³¹

Here the emphasis has been placed on the very possibility of interpretation. Perhaps that is the key to Davidson's appeal to radical interpretation.

³¹ Davidson, "The Structure and Content of Truth", p.325.

From what Davidson has said so far about radical interpretation, it is true that we could infer certain things about the mental states of two speakers each of whom is able to interpret the other successfully and radically. If speaker A is able successfully to interpret speaker B, not only must both A's and B's sets of beliefs themselves abide by, and be founded upon, normative principles of rationality and consistency, but A's and B's principles must be very similar. They could not be too different, or else interpretation of one by the other would not be possible. There must also be massive agreement among A's and B's beliefs. And so if we had reason to believe that for any mindful being at all, any other mindful being could, in principle, successfully interpret the first being on the basis of his potential behavior (i.e., his behavior under all possible circumstances), then we would have reason to conclude that all mindful beings in large part share sets of beliefs and principles of rationality. That would be something we could infer from what Davidson has shown us are the conditions of successful interpretation, and that would be quite a conclusion to be able to draw.

Do we have reason to believe that? More importantly, do we have reason to believe the weaker claim that for any mindful being at all, there is some possible interpreter who could successfully interpret him on the basis of his potential behavior? If this weaker claim were true, we could read off from the conditions upon successful interpretation—from what must be true of a subject's mental make-up for that person to be successfully interpreted—conclusions about the nature of the mental in general, conclusions about the nature of all possible thought.

Of course, Davidson himself clearly endorses this claim. He says that “[w]hat a fully informed interpreter could learn about what a speaker [believes] is all there is to learn,”³² and that “the nature of language and thought is such as to make them interpretable.”³³ And elsewhere he writes, “Thoughts, desires, and other attitudes are in their nature states we are

³² Davidson, "A Coherence Theory of Truth and Knowledge", p.315.

equipped to interpret; what we could not interpret is not thought.”³⁴ A similar idea is found in Quine, according to whom “[t]here are no meanings, nor likenesses nor distinctions of meaning, beyond what are implicit in people’s dispositions to overt behavior.”³⁵ The same is to go for all the attitudes.

Is successful interpretation of a thinker always possible? We obviously have good reason to think that interpretation of thought and belief is possible in one sense. We know it is possible, because we know it often goes on, even if it is not always—or ever—radical. In that sense of possibility, interpretation is indeed possible, and any theory of belief should account for that possibility. From that fact, we can conclude that *some* thought must satisfy those conditions that Davidson has shown must obtain for interpretation to be successful. But from the fact that interpretation is possible in this sense, we cannot conclude that all thought must fulfill those conditions that a particular thinker’s mental make-up must fulfill in order for that thinker to be interpretable. We know it is possible for chairs to be comfortable, for some certainly are. And if it is a condition on a chair's being comfortable that it feels good to sit in, then we can conclude that some chairs feel good. But it would be a mistake to infer from the possibility of chairs being comfortable that it is therefore an essential mark of chair-hood that they feel good.

That is why Davidson could not merely be asking about what is required to make sense of the fact that we ourselves engage in practices in which we do in fact understand each other. From what is required of our thought for us to engage in such practices, Davidson could not legitimately draw conclusions about all possible thought, i.e., about the nature of thought in

³³ Davidson, "The Myth of the Subjective", p.166.

³⁴ D. Davidson, "Representation and Interpretation", in W. H. Newton-Smith and K. V. Wilkes (eds.), *Modelling the Mind* (Oxford: Oxford University Press, 1990), p.14.

³⁵ W. V. O. Quine, "Ontological Relativity", in his *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1977), p.29.

general. Yet the conclusions he draws—for instance, that thought is only possible for creatures in communication with others—seem to concern precisely that.

Davidson would need a stronger premise about the possibility of interpretation than simply that it is possible in the sense above. He needs what I just referred to as the “weaker claim”:

(N) For any thoughtful being, there is some possible interpreter such that, had she access to all of the being’s potential behavior, she could ascertain his beliefs.

I refer to this premise as “(N)”, because it holds that the possibility of interpretation (in the sense in (N)) is *necessary* for thought. If, for any being with beliefs, there is always at least *some* possible being who could in principle interpret that being, and if considerations of radical interpretation illustrate that certain things must be true about a thoughtful being for him to be able to be successfully interpreted, then radical interpretation would help guarantee that those things that must be true about a thinker for him to be able to be successfully interpreted must be true of all thinkers. Every believer would be in principle interpretable. Conclusions about belief *in general* could legitimately be drawn. But if interpretation is not always possible in this sense, i.e., if it is possible for there to be a thoughtful being whom no possible interpreter could interpret, then we would not be justified in concluding that the characteristics with which we endow the mental *in order to make interpretation possible* should be essential to thought, or that all thought must exhibit such properties.

Such a conception of the mental—(N)—does not result from anything that has been said so far about the situation of radical interpretation. Our considerations of radical interpretation have primarily concerned the necessary conditions for successful interpretation, not the range of believers for whom such interpretation is possible. For radical interpretation to acquire its relevance from the possibility of interpretation, this view of the interpretability of the mental must be secured, or at least assumed, prior to the very appeal to interpretation.

(N), however, is a substantial premise, and there are many who would deny it. Colin McGinn, for example, writes,

It is a *condition of interpretability* that the subject by and large believes what he perceives. . . . (This is not to say [that a person who systematically and globally refuses to let his beliefs be shaped by his experience] is *impossible*; it is just that he is not *interpretable*.)³⁶

John Searle would also reject (N). Searle denies that “there is some sort of conceptual or logical connection between conscious mental phenomena and external behavior.”³⁷ He says,

*Ontologically speaking, behavior, functional role, and causal relations are irrelevant to the existence of conscious mental phenomena. Epistemically, we do learn about other people’s conscious mental states in part from their behavior. Causally, consciousness serves to mediate the causal relations between input stimuli and output behavior; and from an evolutionary point of view, the conscious mind functions causally to control behavior. But ontologically speaking, the phenomena in question can exist completely and have all of their essential properties independent of any behavioral output.*³⁸

Elsewhere, when discussing meaning, Searle claims that it is false to suppose that “what isn’t conclusively testable by third-person means isn’t actual.”³⁹ Presumably, he would say the same for belief and the other attitudes. Those who support premises like (N), Searle would say, do not respect the “first-person, ‘subjective’ point of view.”⁴⁰

³⁶ C. McGinn, "Radical Interpretation and Epistemology", in E. LePore (ed.) *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson* (Oxford: Basil Blackwell, 1986), p.167.

³⁷ J. R. Searle, *Rediscovery of the Mind* (Cambridge, MA: MIT Press, 1992), p.69.

³⁸ *Ibid.*

³⁹ J. R. Searle, "Indeterminacy, Empiricism, and the First-Person", *Journal of Philosophy* 84 (1987), p.146.

⁴⁰ *Ibid.*: p.145.

In the remainder of this paper, I want to explore whether there is good reason to endorse (N). It is important to note first, though, that even equipped with (N), Davidson's appeal to the radical interpreter would still yield only *some* of the aspects of belief he thinks can be gleaned from radical interpretation. Even if belief is in principle interpretable, in the sense above, it would still not follow from what Davidson says about the conditions of successful interpretation that constitutive interpretationism is true, or that having a belief does not involve mental objects or representations that the mind somehow grasps. Nor would it follow from Davidson's considerations about interpretation that it is sometimes an indeterminate matter what beliefs a subject has.

One thing that *would* follow, however, is the essentially normative character of belief and thought. If for any thoughtful being there is some possible interpreter who could successfully interpret him, and if successful interpretation requires that the thoughts of the interpreted adhere to normative principles of rationality and consistency, then the thoughts of all possible thinkers must satisfy such principles. Likewise, if for interpretation to be successful an interpreter must take her subject's most basic beliefs to be about the very things in the world that cause those beliefs, and if successful interpretation is always possible, then the most basic beliefs of all possible thinkers must generally be about the very things in the world that cause those beliefs. And indeed, it is this latter conclusion concerning the connection between cause and content that provides the basis for Davidson's argument against skepticism of the senses as well as for one of his arguments for the social character of thought.

Before looking at the explicit support Davidson himself provides for (N), it is worth addressing one plausible idea that might appear to lead to (N). One reason someone might be inclined to accept (N) is that what human beings do—how they behave—is in part explained by their mental states. I take an umbrella with me when I leave the house, because I believe it might rain (and because I desire to stay dry, etc.). Mental states and events explain what we do.

Perhaps for any particular action most of my mental states will not factor into an explanation of it. My belief that Van Gogh was a painter does not help to explain my taking an umbrella. What may be true, though, is that for each of my mental states, there are possible circumstances under which the behavior I would engage in would be partly explained by that state. About intention, for instance, Charles Taylor writes,

This is part of what we mean by ‘intending X ’, that, in the absence of interfering factors, it is followed by doing X . I could not be said to intend X if, even with no obstacles or other countervailing factors, I still didn’t do it.⁴¹

The same would go perhaps (though not as directly) for all propositional attitudes.

Someone might embrace (N), then, because she endorses some premise like the following:

(C) For every one of a subject’s mental states, there are (perhaps infinite) possible circumstances under which that mental state would play a causal role in bringing about some piece of behavior.

Some readers will be less comfortable with (C), which concerns causation, than with a premise that emphasizes explanation, such as:

(C*) For every one of a subject’s mental states, there are (perhaps infinite) possible circumstances under which that mental state would serve to explain some piece of behavior.

But let us grant (C) for the sake of argument. (C*) cannot support (N) for the same reasons that we will see that (C) cannot support (N).

If (C) is correct, then for every one of a subject’s beliefs there would be some possible circumstances in which that belief would play a causal role in bringing about some piece of behavior, and an interpreter would in principle have access to all of those circumstances and

⁴¹ C. Taylor, *The Explanation of Behaviour* (London: Routledge & Kegan Paul, 1964), 33.

behavior. However, for any piece of behavior, there are countless ways in which it could come about, countless combinations of mental states that could cause one to behave in that way. Davidson's discussion of the connection between meaning and belief reveals the point well. One holds a particular sentence to be true (and thus assents to that sentence) both because of what one means by the sentence and because of what one believes. The attitude of holding true, we remember, is a "vector" of meaning and belief. And so for any sentence, there are many (perhaps infinite) combinations of meaning and belief that could be the cause of someone's assent to the sentence. That itself does not imply that an interpreter could not determine which of all those possible combinations was the one the subject had. Often interpreters do ascertain this. The question for our purposes is whether one always could. Indeed, I think we can imagine thinkers for whom it is plausible to claim that there is no such interpreter. They are admittedly far-fetched cases, not ones we normally confront. However, because Davidson's conclusions are to apply to all *possible* belief, it is appropriate for us to appeal to such cases. One of Davidson's conclusions, we remember, is that thought is possible only for linguistic creatures.

Consider, then, a person whose overriding goal in life is to deceive those who attempt to identify his thoughts. For whatever reason, this goal is of such high priority to him that he would sacrifice his life in order not to reveal his mental make-up. There strikes me as little *prima facie* reason to suppose either that such a person is not possible, or that there is a possible interpreter who would be able to determine this person's mental make-up from his potential behavior. It is not clear, for instance, that any interpreter would be able to discern *why* the deceitful person is so intent on deceiving, or even *that* he is intent on deceiving. The possibility of this thinker is not incompatible with (C). The subject's behavior would still be a causal effect of his mental states. His behavior is the result of his (strange) intentions (as well as of his

beliefs, desires, etc.) Or consider the possibility Colin McGinn mentions in the passage I have already quoted:

It is a *condition of interpretability* that the subject by and large believes what he perceives. . . . (This is not to say [that a person who systematically and globally refuses to let his beliefs be shaped by his experience] is *impossible*; it is just that he is not *interpretable*.)⁴²

Indeed, it seems we could concoct countless such examples.

It might be protested that these sorts of thinkers are irrational. But our countenancing irrational thought is not inappropriate at this juncture. The possibility of irrational thought or action has not yet been ruled out. Davidson may hold that there is a limit upon how much irrationality an interpreter can find in her subject. But to suppose that irrationality—even a hearty dose—is not possible because an interpreter would not be able to determine which mental states her subject has (or even whether he has any) would require the presupposition that successful interpretation is always possible, or that there is no more to the mental than what a fully informed interpreter could ascertain from all of a subject’s potential behavior. And that is precisely the view under consideration. The constraint of rationality we have granted so far concerns only what must be the case if successful interpretation is to be possible.

Let us turn now to the way in which Davidson himself supports theses such as (N). He introduces the situation of radical interpretation in “A Coherence Theory of Truth and Knowledge” in this way:

A speaker who wishes his words to be understood cannot systematically deceive his would-be interpreters about when he assents to sentences—that is, holds them true.

As a matter of principle, then, meaning, and by its connection with meaning, belief

⁴² McGinn, "Radical Interpretation and Epistemology", p.367.

also, are open to public determination. *I shall take advantage of this fact in what follows and adopt the stance of a radical interpreter when asking about the nature of belief.* What a fully informed interpreter could learn about what a speaker means is all there is to learn; the same goes for what the speaker believes.⁴³

In the italicized sentence, Davidson appears to acknowledge that the publicity of meaning and belief is what *allows* him to appeal to the situation of the radical interpreter. Were belief not known to be public, he intimates, the relevance or justification for adopting that “stance” might not be at hand.

But the little support Davidson gives for this now prior claim about the publicity of belief is itself unclear. Davidson appeals to the fact that meaning is public and that so too must be belief, because of belief’s “connection with meaning.” But when briefly discussing how it is that belief depends upon meaning a few paragraphs above, all Davidson says is that “[b]elief, however, depends equally on meaning, for the only access to the fine structure and individuation of beliefs is through the sentences speakers and interpreters of speakers use to express and describe beliefs.”⁴⁴ This is the insight that in order successfully to attribute beliefs to a given speaker, one must know something about what that speaker means by his words. But even if that were true, that dependence, or “connection,” is thus far only an epistemological one. It is a connection that must exist and that the interpreter must rely upon if the project of belief or meaning attribution is to be successful. But there is nothing about that epistemological connection that reveals that beliefs cannot exist where meanings do not, or that they are essentially interpretable, even if meanings are so. Or at least there isn’t without some prior view of belief already in place.

⁴³ Davidson, "A Coherence Theory of Truth and Knowledge", p.315, italics mine.

⁴⁴ Davidson, "A Coherence Theory of Truth and Knowledge", pp.314-315.

Even if we grant Davidson a conception of *meaning* as essentially interpretable, it still would not be clear that he would be justified in concluding anything general about the nature of belief.⁴⁵ As we have seen, in order to determine a speaker's meanings, one must know something about his beliefs. So if the meanings of all speakers are interpretable, then perhaps so are their beliefs.⁴⁶ But that would be to say something only about the beliefs of beings who are involved in communication; it would not be to say anything about all possible belief, or about the essential nature of belief.

Davidson might be justified in drawing a conclusion about belief in general had he already secured the conclusion that only linguistic beings, or those involved in communication, are capable of thought. If the beliefs of linguistic beings are always in principle interpretable, and if only linguistic beings are capable of thought, then the beliefs of all beings with belief would be in principle interpretable. And indeed, Davidson does claim that only linguistic beings are capable of thought. Perhaps then, one might suppose, it is Davidson's reasoning for *this* claim that is what secures (N). But a close look at Davidson's reasoning for this claim, to which I now turn, reveals that it too depends upon (N) (or on something even stronger) and thus cannot support (N).

There are two, perhaps related, routes by which Davidson arrives at this conclusion about the impossibility of thought without language. One emphasizes the importance and origin of the very *concept of* belief; the other emphasizes the necessity of the process of triangulation.⁴⁷ It is

⁴⁵ Of course, one who held that meaning is essentially public might hold that belief is public for the very same reason, i.e., not because one holds that meaning is public and that belief is connected to meaning, but because the very considerations that lead one to conclude that meaning is public might equally apply to belief. But that is not how Davidson here explains his reasoning.

⁴⁶ Even this may not be true. What follows from the fact that all of one's meanings are interpretable may be only that some of one's beliefs are interpretable.

⁴⁷ The first is found, e.g., in Davidson, "Thought and Talk", D. Davidson, "Rational Animals", in E. LePore (ed.), *Actions and Events: Perspectives on the Philosophy of Donald Davidson* (Oxford: Basil Blackwell, 1985), and Davidson, "Three Varieties of Knowledge". The second is found, e.g., in Davidson, "Three Varieties of Knowledge", and Davidson, "Epistemology Externalized".

not clear whether Davidson considers these arguments entirely distinct, or whether the latter is something of a development of the former. I will consider them each separately.

In order for a creature to have a belief at all, Davidson argues in the first of those arguments, that creature must have the *concept* of belief. Having that concept involves grasping the distinction between truth and error, understanding that there is a difference between something's being the case and something's only seeming to be the case, having the concept of objectivity. The only way in which a creature could ever attain that concept (or concepts), Davidson believes, is through interpersonal communication, through the context of interpretation. We remember that we "come to understand" the concept of belief just in this connection, and that, as he says, "We have the idea of belief only from the role of belief in the interpretation of language. . . ." ⁴⁸ Therefore, any creature capable of thought must be, or must have been at some point, involved in communication with others. Those two steps inform Davidson's first argument for thought's dependence upon language.

Were we to grant Davidson his contestable first premise (that for one to have a belief at all one must also have the concept of belief), Davidson's reasons for regarding the context of interpretation as being the only situation in which one could ever attain that concept are still scarce and difficult to discern. In one of the few places where he appears to give any argumentation at all for this claim, Davidson alludes briefly to what he takes to be Wittgenstein's private language argument.⁴⁹ However, there is little explanation of the argument nor any discussion of how it is to be extended to his particular point about the conditions for one's having the concept of objectivity. Davidson claims that having this concept requires having "the standard provided by a shared language," but he does not explain why that is.

⁴⁸ Davidson, "Thought and Talk", p.170.

⁴⁹ Davidson, "Three Varieties of Knowledge", p.157.

Besides that, however, there is little else given to convince potential objectors that the concept of belief could not possibly be acquired in some other way, for example, innately, or that one's acquiring the concept is not simply a matter of acquiring a particular brain state, as some philosophers would have it. Davidson himself admits in "Rational Animals" that "[t]o complete the 'argument,' however, I need to show that the *only* way one could come to have the subjective-objective contrast is through having the concept of intersubjective truth. I confess I do not know how to show this. Neither do I have any idea how else one could arrive at the concept of an objective truth."⁵⁰

Davidson may be right that the concept of belief is necessary for successful interpretation, and thus that any successful interpreter must have it. But that does not imply that the context of interpretation is where the interpreter acquires that concept. Of course, for one who already holds that belief and thought (and therefore concepts too) arise only in contexts of interpretation, it will naturally follow that the particular concept of belief can only be had by those involved in interpretation. But that would be to presuppose the very thesis about the social character of thought for which we are seeking support.

Of course, one *might* get the concept from the situation of interpretation. Davidson provides a plausible, perhaps enticing, suggestion as to where we *do* get this concept. But Davidson gives little reason to suppose, not just that the way he suggests we acquire the concept is the way that we definitely *do* get it, but that it is the way that we, or any creature at all, *must* get it. Yet that latter claim is what is required.

Similar problems afflict Davidson's second argument. That argument appeals to the way in which our beliefs and meanings acquire their content. In the most basic cases, Davidson argues, the contents of our beliefs are determined by the objects and events in the world that cause

⁵⁰ Davidson, "Rational Animals", p.480.

them; those causes are what those beliefs are in general about. But for any given belief, as for any event at all, there are many causal chains extending back in time. Every one of those chains consists of many events each of which might appropriately be considered a cause of that belief. (The Big Bang, for example, Davidson points out, is one of the many causes of any given belief.) What is necessary, then, to determine the *unique* cause that is what determines the content of a particular belief and that is what that belief is therefore about, Davidson believes, is a process of triangulation that occurs between two or more people. Triangulation is necessary for beliefs to have content, he argues, for only that process could determine, of the many causes, the unique cause that gives a particular belief its content. Belief is thus essentially social, as there could be no content at all apart from the context of triangulation.

However, Davidson supports the first premise of this argument—that the contents of beliefs are in the simplest cases determined by those things that cause those beliefs—once again only by appeal to radical interpretation, by looking at how a radical interpreter would have to go about interpreting her subject's beliefs. What an interpreter would take her speaker's words to refer to are those objects or events in the environment that the interpreter thinks systematically cause the speaker to utter those words (e.g., rabbits).⁵¹ That is why Davidson believes that “we can't in general first identify beliefs and meanings and then ask what caused them.”⁵² And that seems correct. But these are still just facts about the successful attribution of belief. The purported fact that an interpreter must take her subject's beliefs to be about those things that she believes cause them, and that those beliefs must really be about such things, does not imply anything about what determines the contents of *any* mindful being's beliefs, or about what is necessary for a belief to have content at all or for a mindful being to have belief at all. The

⁵¹ The sorts of interpretation Davidson considers when he argues for this premise are not always “radical.” For example, see Davidson, “Epistemology Externalized”, pp.194-195.

⁵² Davidson, “A Coherence Theory of Truth and Knowledge”, p.317.

inevitability of successful interpretation is precisely what is in question; the relevance of investigating the conditions on successful interpretation for understanding the nature of belief in general is still the subject of the larger inquiry of which our discussion about the social nature of thought is merely a part.

These arguments for the social character of thought, then, cannot provide the required support for the thesis that belief is in principle interpretable. The success of those arguments still depends upon one or another significant presupposition about belief which has yet to be secured. Indeed, the most likely candidate to provide the support for those arguments is the very presupposition we have been looking to Davidson's arguments for the social character of thought to *support*: the thesis that belief is in principle interpretable. Those arguments cannot support that thesis, because they depend upon it. And his reasoning for that thesis, we have seen, seems to rely on the conclusion of those arguments: the social character of thought. They both involve a prior confidence in the relevance of radical interpretation for understanding the nature of the mind. Indeed, it is interesting that even those arguments for the social nature of thought require such relevance, because sometimes Davidson seems to intimate that one reason we may feel justified looking to the radical interpreter is precisely that thought is only possible for those in communication.⁵³

5. Conclusion

We are therefore left with no reason from Davidson's philosophy to suppose that (N) is true. This is of importance given that Davidson's appeal to the situation of radical interpretation for the purpose of securing his substantial conclusions about mental phenomena depends precisely upon some premise like (N). Indeed, once we appreciate that dependence, it becomes

⁵³ Davidson, "Three Varieties of Knowledge", p.157.

much less surprising that Davidson is able to draw such formidable conclusions about the mental merely from considerations about the conditions upon attribution. The bulk of the work is being done by an unsupported presupposition.

That Davidson's endorsement of a substantial conception of the mind comes prior to his appeal to radical interpretation explains the parallel with which I began, between the method of Davidson's philosophy and what Davidson's philosophy is about or purports to show. The mind emerges from contexts of interpretation, just as Davidson's conclusions about the mind emerge from his inquiries into interpretation. Both strands of the parallel are grounded in a prior conviction that the situation of interpretation is relevant for understanding the nature of the mind. Such a conviction goes hand-in-hand with the view of the mental as in principle interpretable. Only such a view as (N) (or something stronger, such as the constitutive view) could make successful an appeal to interpretation for many of the ends Davidson asks of it. Whether Davidson's conviction that radical interpretation is relevant for understanding the mind stems from a prior subscription to (N) (or to the constitutive view), or whether his subscription to (N) stems from a prior conviction that radical interpretation is relevant, is unclear.

Of course nothing I have said in this paper shows that any of these prior views is false. Nor does anything I have said show that any of the conclusions Davidson draws about the mental from investigating the situation of radical interpretation is false. Indeed, those who are partial to one of the required presuppositions may well find Davidson's appeal to radical interpretation to bear considerable promise. From what is learned about the conditions upon successful interpretation—about what a subject's thought must be like in order for him to be successfully interpreted—we could draw conclusions about all possible thought, i.e., about the nature of thought in general.

Whether those who endorse (N) would agree with the *particular* conclusions Davidson draws about the nature of the mental would depend upon whether they agree with the particular

conditions Davidson identifies. One still could not abstract anything like the constitutive view of the mental from a consideration of radical interpretation. That is precisely the sort of view one might embrace prior to one's appeal to radical interpretation. It entails (N). If, however, one agrees with Davidson that a condition upon successful interpretation is that a subject's mental make-up abides largely by normative constraints of rationality and consistency, and one embraces (N), then one might agree that the mental must abide by normative constraints of rationality (and also, in turn, perhaps agree with Davidson's claim about the anomalism of the mental, even anomalous monism). If one agrees with what Davidson concludes about the relation in which the contents of one's beliefs must stand to their causes in order for there to be successful interpretation, and one embraces (N), then one might endorse Davidson's refutation of skepticism. And if one also agrees with the purported fact that some process of triangulation is what determines precisely which causes stand in that relation to one's beliefs, then one might also go along with Davidson's conclusion that thought is only possible for linguistic beings.

But for any of these lines of argument to be successful, one would need to provide grounds for believing (N). Those grounds are not to come from the situation of radical interpretation, as Davidson suggests, but must come from elsewhere.⁵⁴

What are we to conclude from the fact that there may be such a circularity in Davidson's philosophy? Did Davidson simply fail to see this? Or did he see it but not consider it a handicap? Not all circularities are vicious. Indeed, some philosophers have argued that transcendental philosophy itself (of which Davidson's philosophy is often considered a form) is inherently circular, yet that the circular nature of transcendental philosophy does not constitute a

⁵⁴ William Child, who also emphasizes the importance of (N) for many of Davidson's arguments concerning the mental, finds compelling support for it in the work of Ludwig Wittgenstein (Child, *Causality, Interpretation, and the Mind*, pp.35-37). Whether Wittgenstein's philosophy provides good reason for believing (N) is unfortunately not something I can pursue here.

problem or difficulty for it.⁵⁵ I suspect that the circularity I have identified is problematic, that it reveals that Davidson's work does not contain adequate justification for the controversial view of interpretation upon which many of his arguments depend. At the very least, a great deal would certainly need to be said in order to show that the view is justified by way of this circle. I do not have the space here to investigate the prospects of showing this. What I hope to have done here is to have established that such a circularity does exist in Davidson's philosophy.⁵⁶

⁵⁵ J. Malpas, "The Transcendental Circle", *Australasian Journal of Philosophy*, 75 (1997), pp.1-20. The sort of circle Malpas identifies is different from the one I have identified.

⁵⁶ I am grateful to Jason Bridges, Donald Davidson, John Heil, Sean Kelly, Jeff Malpas, John Searle, Barry Stroud, Bruce Vermazen, and Wai-hung Wong for their help on earlier drafts of this paper.