

## DAVIDSON'S EXTERNALISM AND THE UNINTELLIGIBILITY OF MASSIVE ERROR

Andrew Carpenter

### 1. Global scepticism and its refutation

In *The Significance of Philosophical Scepticism*, Barry Stroud considers a haunting question: "Could not the external world be completely different from what we perceive it and believe it to be?" (Stroud, 1986: 210). Even the least persuasive sceptical arguments can evoke this doubt. For example, although the conclusion that our perceptual beliefs are in general untrustworthy certainly doesn't follow, directly at least, from the fact that our senses sometimes lead us astray, reflecting on instances of illusion and perceptual error can engender the arresting thought that we might be experiencing systematic and general error or illusion. From this thought — the thought that our empirical knowledge-claims might generally be false or at least insufficiently justified — sceptical arguments are not far behind. To give one prominent argument: if knowledge is closed under known deduction, then, if we have empirical knowledge, we must know that we aren't in massive error; since global scepticism is false if we have any empirical knowledge, the closure principle entails that we can have empirical knowledge only if we know that the global sceptical possibility does not obtain. The trouble, of course, is that it is unclear how we could know this about the global sceptical possibility, especially, as seems reasonable, if appealing to putative instances of empirical knowledge is ruled out as question-begging.

The closure argument tells against the policy of simply refusing to take seriously global sceptical doubts. But how might such doubts be dealt with? First note that, by itself, the sceptical possibility has no obvious direct epistemic force: drawing sceptical conclusions requires something else, viz. an argument to parlay the doubts into conclusions about the difficulty of justifying empirical beliefs. Consequently, one way of refuting global scepti-

*Disputatio* 4 (1998)

cism would be to block this move. In the light of the number of arguments advanced for global scepticism, however, this project quickly attains Herculean proportions. Another tactic would be to engage the sceptical doubt itself, by arguing against the intelligibility of the very thought that the external world could be completely different from how we believe it to be. Since all versions of global scepticism arise from this thought, an argument holding that such massive error is incoherent or unintelligible would show that all arguments for global scepticism must be unsound. Note also that if the original sceptical question is unintelligible, then it will follow that a goal which has been central to traditional epistemology, attempting to answer that question from some “transcendental viewpoint”, is superfluous or perhaps even incoherent. Understanding the form and scope that “naturalised epistemology” might take once this goal is forsworn is an important task; in this essay, however, my discussion will be limited to the project of showing that the sceptical question is unintelligible. My specific target shall be an anti-sceptical argument of this type offered by Donald Davidson.

Before considering Davidson’s arguments, it will be useful to see why this anti-sceptical strategy is unavailable to anyone who accepts a characteristic type of empiricist epistemology. There are many causal intermediaries along the causal path from external objects to our empirical beliefs. It is easy to see that the global sceptical possibility is coherent under any “epistemological intermediary” theory, according to which the ultimate source of justification or meaning is some causal intermediary short of the very external objects we take our thoughts to be about. Epistemological intermediary theories are just those theories that take some intermediate cause — experience, sense data, the given, sensory evidence, etc. — as the source of justification or meaning. Obviously, it is always a salient question what are the causal antecedents of an epistemic intermediary — that is, what external objects or events are earlier in the causal sequence. Thus global sceptical doubts must be admitted as intelligible if epistemic intermediaries are countenanced, since we are completely mistaken in our beliefs about the world just if the causal antecedents of the epistemic intermediaries are not what we take them to be. In other words, it appears that the anti-sceptical strategy of showing the unintelligibility of massive error can succeed only if epistemic intermediaries are eschewed and external objects themselves are held to be the ultimate source of justification and meaning.

If my argument so far is correct, Davidson's general anti-sceptical strategy, which attempts to demonstrate the unintelligibility of global sceptical doubts, should be understood in the context of (and as presupposing) his rejection of epistemic intermediary theories. In opposition to the "proximal" and "internalist" positions built upon the so-called third dogma of empiricism, Davidson's anti-sceptical argument rises from a distal, externalist theory of meaning. Externalism roughly is the doctrine that mental states are supervenient on (among other things) external objects. As we shall see, an important consequence of this doctrine is that the content of our propositional attitudes is not independent of, but is (in part) determined by, the very external objects they purport to be about.

We have seen that the global sceptical possibility can be cashed out in terms of a "causal gap" between our thoughts and the world just because it seems that radically different causes could produce the very same beliefs about the world we currently possess. Radical sceptical doubts can be generated by holding our mental states constant and varying the external objects to which they are causally connected — that is, by claiming that our current mental states could have had various, radically different causal antecedents. By trading on the idea that our sentences could mean the very same thing even if the world were radically different, this move obviously presupposes the "internalist" view that there is no necessary connection between meaning and what is true of the outside world. By contrast, externalism rejects the legitimacy of "holding constant" a person's mental states while varying his external environment; externalism and internalism differ precisely on whether the individuation of propositional attitudes is independent of their causal antecedents. Interpreted in the light of externalism, the significance of the alleged "causal gap" is this: the doctrine of the supervenience of mental states on external objects entails that if different causal antecedents had the same physical effects on us, then people in the different environments would have different thoughts. Despite (by hypothesis) undergoing the same physical stimulation, different thoughts must be attributed to people in the various situations precisely because, on the externalist hypothesis, there is a necessary connection between the content of our thoughts and their actual causal antecedents.

## 2. A knockdown argument against scepticism?

Davidson shares with Kant an extremely powerful anti-sceptical strategy. Both philosophers argue that global scepticism about the external world must be false because its truth would violate a condition necessary for belief and thought. If successfully deployed, this strategy would leave the sceptic no room to manoeuvre. On the usual formulation at least, the global sceptical doubt does not question that we have thought or belief, but rather plants the seeds of suspicion that all of our beliefs about the world could be false. Indeed, it makes no sense to raise the sceptical question itself — “Could not the external world be completely different from what we perceive it and believe it to be?” — unless the questioner takes it for granted that the conditions necessary for thought are met. It follows that there is a quick answer to the question of whether a sceptic might be prepared to bite the bullet and actually endorse the impossibility of thought; even if it could somehow be presented in a way that isn’t straightforwardly self-refuting, that position, which seems more nihilistic than sceptical, isn’t the intended target of arguments which attempt to answer what I’ve called the sceptical question. The target of this paper is Davidson’s proposed answer to that question.

The claim that Davidson’s anti-sceptical strategy is similar to Kant’s may sound bizarre. After all, although he argued that only one scheme is possible for us, the fundamental place Kant accorded the distinction between concepts and intuitions makes him a philosopher of scheme and content *par excellence*. Yet the affinities are striking. Kant attempts to draw ontologically-significant conclusions from reflections on the conditions necessary for judgement. Davidson does something quite similar, viz. he attempts to expose the ontological import of certain conditions necessary for belief. In the light of the intimate connection in Kant’s philosophy between judgement and belief, these projects converge; both philosophers aim at a refutation of global scepticism of the strong type sketched above by trying to show why, necessarily, our empirical beliefs must generally be true. Of course, Davidson’s theory of meaning differs drastically from Kant’s. One consequence of this difference is that whereas Kant’s arguments take the possibility of experience as their starting point, Davidson’s point of departure is the possibility of communication.

A long list of similarities between Kant’s and Davidson’s arguments notwithstanding, the profound differences between Kant’s and Davidson’s

positions make it unlikely that Davidson's arguments will be vulnerable to many of the criticisms which can be levelled at Kant's. However, if I am correct in holding that Kant and Davidson share a basic anti-sceptical strategy, any deep problems in Kant's attempt to carry out this strategy might prove sticking points for Davidson's argument as well. There is, I think, a fatal flaw at the heart of Kant's anti-sceptical argument. On my reading Kant's argument requires a coherence theory of both knowledge and truth; as I understand it, the argument in the Refutation is invalid unless a coherence theory of truth is assumed. The intuition that any argument for a coherence theory of truth will somewhere beg the question against the Cartesian sceptic has, to my mind at least, some plausibility. If, as I believe is possible, this prediction can be defended, Kant's commitment to a coherence theory of truth undermines his attempt to carry out his extremely ambitious anti-sceptical strategy.<sup>1</sup>

---

<sup>1</sup> The interpretation and criticism of Kant sketched here are developed in my B.Phil. thesis, "Transcendental Arguments and Transcendental Idealism". Despite the failure of Kant's own arguments, in that work I argue against the pessimistic conclusion that it is impossible to construct a successful Kantian anti-sceptical transcendental argument. I should note that, although I briefly compared Davidson and Kant in my thesis, Carol Rovane's "The Metaphysics of Interpretation" prompted me to explore the connections more fully.

Following Ralph Walker's interpretation of Davidson in *The Coherence Theory of Truth*, I had initially suspected that Davidson might be susceptible to this line of criticism. However, despite the title of his essay "A Coherence Theory of Truth and Knowledge", on reflection it seems that Davidson is committed to a coherence theory of justification but not to a coherence theory of the nature of truth. In recent "Afterthoughts", Davidson cites two reasons why the title of his essay is misleading. He explains, first, that "my emphasis on coherence was... just a way of making a negative point, that all that counts as evidence or justification for a belief must come from the same totality of belief to which it belongs" (1987:135). Although this negative point is consistent with a coherence theory of truth, it is only committed to a coherence theory of justification; Davidson comments that "this... claim has typically led those philosopher who hold it to conclude that reality and truth are constructs of thought; but it does not lead me to this conclusion" (*ibid.*). Second, Davidson holds that, because he offers "a Tarski-style theory of truth", it is wrong to understand him "as providing an *explanation* or *analysis* of truth" (*ibid.*). Constructing a theory of the nature of truth isn't part of Davidson's project because, on his view, "truth is as clear and basic a concept a we have... [and] and any further attempt to explain, define, analyse, or explicate the concept will be empty or wrong" (*ibid.*).

I have claimed that Davidson shares Kant's ambitious anti-sceptical strategy: like Kant, he attempts to show that belief has anti-sceptical necessary conditions. If it can be carried out, this strategy would provide a decisive argument against global scepticism. The crucial hurdle is that the argument for the claim that belief carries with it anti-sceptical presuppositions must not beg the question; if a non question-begging analysis of belief cannot be constructed, then, obviously, the strategy cannot be as powerful as I have maintained. I have suggested that Kant's argument stumbles at just this point. A main objective in what follows will be investigating whether Davidson is better able to realise the potential of the project he shares with his illustrious predecessor.

### 3. Radical interpretation

My ultimate target is Davidson's so-called "omniscient interpreter" argument and the externalism lying behind it. Davidson gives this argument in several articles and, *prima facie* at least, there appear to be two quite different versions, viz. an earlier version in "The Method of Truth in Metaphysics" and a later version in "A Coherence Theory of Truth and Knowledge". Although the earlier argument has attracted much critical attention, here I shall discuss the argument from the more recent essay. My reason for doing this is simply that in this argument, but not explicitly in the earlier one, externalism plays a prominent role.<sup>2</sup>

The omniscient interpreter argument falls out of Davidson's account of radical interpretation. Radical interpretation, in turn, rests on Quinean methodological presumptions about theories of meaning. Like Quine, Davidson accepts a "third-person" or behaviouristic approach to meaning. This position is based on what looks like a truism. It seems obvious that our language bears important inferential relations to the content of our thoughts; in short, all meaning is linguistic meaning. Since language can be conceived as a public institution, an attribute of linguistic meaning is inter-subjectivity. The Quinean claim follows that a theory of meaning must be behaviouristic in a relatively mild, non-reductionist sense: linguistic meaning, and therefore the

---

<sup>2</sup> To what extent the two formulations of the omniscient interpreter argument represent a change in Davidson's thought is an interesting question; however, it is perhaps best to stipulate that topic as falling outside the scope of this essay.

content of all our propositional attitudes, must in principle be derivable from observable behaviour.

In this essay I shall not examine any arguments for the claim that meaning must be public; I shall proceed by taking this for granted and investigating whether any anti-sceptical claims can be developed from it. I don't think that assuming the inter-subjectivity of meaning begs the question against traditional scepticism about the external world. It seems a coherent thought that, even if meaning must be public in the way assumed, the external world could be completely different from how we take it to be; if the Quinean attitude toward meaning ultimately bears any anti-sceptical fruit, the work will be done in the details of a worked-out third person theory of meaning. In Davidson's case, the anti-sceptical labour occurs in his defence of the externalist claim that the meanings of certain sentences are determined by the circumstances that cause a speaker to hold them true.

An important consequence of the Quinean approach towards meaning is that theories of meaning must be extensional, not intensional.<sup>3</sup> By taking seriously the social character of meaning, this approach takes as its fundamental objective making out meaning on a wholly observational basis, i.e. without taking for granted an antecedent grasp of the meaning of any of a speaker's utterances. The idea is that all meaning is discoverable on the basis of empirical knowledge of extensional relations holding between speakers and their utterances. In Davidsonian terms, the Quinean approach to meaning takes it for granted that meaning is discoverable through a process of public interpretation. Davidson's account of radical interpretation is a description of how one can get from extensional facts about which sentences a person "holds true" to intensional facts about what that person means by her sentences and about what she believes.

Since "holding true" is an extensional relation between people and their sentences, one can know which sentences a person holds true while being wholly ignorant of what that person takes those sentences to mean. It is hard to overemphasise the difficulty of individuating propositional attitudes while

---

<sup>3</sup> Much more can be said about this point than I say here. My intention is to sketch only the barest outlines of radical interpretation necessary for understanding the omniscient interpreter argument. Most steps I summarise without comment, and many others I skip. In the last category there are many minor points (for example the reasons why the sentence is taken as the basic observable unit) and some large ones as well. Most notably, I make no mention of any indeterminacies.

remaining faithful to the constraints of an extensional account of meaning. Consider two apparent truisms. First, given meaning, it is a trivial task to determine the content of someone's beliefs. Second, if you know the meaning of one term, you will possess insight into the meaning of many. On the one hand, if you know a sentence someone holds true and you know what she means by that sentence, then you know the content of one of her beliefs; if you take meaning as given, it is an easy matter to move from the extensional 'A holds *s* true' to the intensional 'A believes that *p*' (assuming that *s* means that *p*). However, since extensionalists can take as given only extensional facts, the problem for the radical interpreter is to reach facts about *both* meaning and belief from an extensional starting point. The lesson to be drawn from the second truism, on the other hand, is that this can be done only by adopting a holistic approach to meaning, viz. by holding that you know the meaning of a sentence only when you know its place in a large network of sentences. If you know the meaning of one of a speaker's sentences, you will be in a position to know the meaning of many other of that speaker's sentences, just as if you know what you mean by 'that is a rabbit' or 'grass is green' then you must know the meaning of a large number of sentences about rabbits, grass, colours, animals, plants, etc. The moral is that just as extensional interpreters must strive to attain both meaning and belief simultaneously, so they are incapable of determining meaning on a sentence by sentence basis; extensionalists are compelled to adopt a holistic approach to meaning.

Much more can be said about both these features of radical interpretation. However, we have seen enough to understand how Davidson can forge a strong connection between having a propositional attitude and the extensional relation of "holding true". In particular, it follows that a doubt can have content only if the doubter holds true many other beliefs. It is impossible to doubt that something is a rabbit without holding true, among others, many beliefs about what a rabbit is. Likewise, the global sceptical doubt has content only if we hold true a host of empirical beliefs. Although I haven't discussed this move yet, at this point I will simply stipulate that, because he applies the principle of charity across the board, Davidson also is in a position to insist that we all must hold true roughly the same "total theory" about the world. It follows from this that there is no way to discover that a speaker is largely wrong about the world.

The question to consider is whether this conclusion possesses any anti-sceptical force. Why shouldn't it be interpreted as showing that we must share a common conception of the world, regardless of whether or not our beliefs about it are in fact largely veridical? Thus far, the argument I have discussed seems to show that massive disagreement, but not massive error, is impossible. After all, "holding true" is not necessarily "holding-truly".

#### **4. The omniscient interpreter and externalism**

The objection is that radical interpretation has no anti-sceptical force because it fails to forge a connection between belief and truth. Davidson would insist that a fundamental point has been overlooked, viz. his externalism. In discussing this objection in "A Coherence Theory of Truth and Knowledge", Davidson appeals to the externalist principle that "we can't in general first identify beliefs and meanings and then ask what caused them... [t]he causality plays an indispensable role in determining the content of what we say and believe" (1983: 131). It is in the context of discussing this claim that Davidson gives the omniscient interpreter argument:

Why couldn't it happen that speaker and interpreter understand one another on the basis of shared but erroneous beliefs? This can, and no doubt often does, happen. But it cannot be the rule. For imagine for a moment an interpreter who is omniscient about the world, and about what does and would cause a speaker to assent to any sentence in his (potentially unlimited) repertoire. The omniscient interpreter, using the same method as the fallible interpreter, finds the fallible speaker largely consistent and correct. By his own standards, of course, but since these are objectively correct, the fallible speaker is seen to be largely correct and consistent by objective standards. We may also, if we want, let the omniscient interpreter turn his attention to the fallible interpreter of the fallible speaker. It turns out that the fallible interpreter can be wrong about some things, but not in general; and so he cannot share universal error with the agent he is interpreting. Once we agree to the general method of interpretation I have sketched, it becomes impossible correctly to hold that anyone could be mostly wrong about how things are. (*Ibid.*)

The first claim I will make about this argument is that appealing to the idea of an omniscient interpreter begs no question against the sceptic; after all, the global sceptical doubt could be characterised in terms of a distinction between lowly, imperfect human beliefs and the necessarily veridical beliefs of an omniscient intelligence. Another potential stumbling-block is the

thought that there is something inconsistent in Davidson's making the omniscient interpreter adopt "the same method as the fallible interpreter". Wouldn't part and parcel of being omniscient include knowledge of everyone's propositional attitudes, thus obviating the need for such a being to interpret? One answer to this question is just that, to be faithful to our extensionalist approach to meaning, we must hold that even omniscient beings have no choice but to adopt a behaviouristic "third-person" attitude towards meaning. A deeper answer is this: there is, on Davidson's view, no objective "fact of the matter" about the content of a person's propositional attitudes, independent of what an interpreter can correctly make out. In Davidson's words, "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"; his position, of course, is that the same goes for the other propositional attitudes as well (1983: 129).

When I alluded to the principle of charity in the previous section, I stipulated my acceptance of Davidson's position that "from the interpreter's point of view... [there is no way] he can discover the speaker to be largely wrong about the world" (1983: 131). Thus it is impossible that an omniscient interpreter could correctly interpret someone as massively mistaken. We turned to the omniscient interpreter argument after considering this question: is there any reason to suppose that Davidson's blanket application of the principle of charity entitles him to conclude that massive error, and not just massive disagreement, is impossible? Since, on the one hand, the omniscient interpreter necessarily "finds the fallible speaker largely consistent and correct" and, on the other, the omniscient interpreter is omniscient and thus his standards are "objectively correct", the conclusion seems to follow that the speaker's beliefs must actually be by and large "objectively correct", i.e. largely true.

The question of whether an omniscient interpreter really exists seems of no consequence to the anti-sceptical argument, which relies on the thought that if our beliefs about the world were generally mistaken, then *per impossibile* an omniscient interpreter, were there one, would correctly interpret us as massively mistaken. There is also a question of whether Davidson's use of the principle of charity begs the question against the sceptic. However, this objection has no force if Davidson can substantiate his claim that successful interpretation requires its application. Since "what a fully informed interpreter could learn about what a speaker means is all there is to learn"

about the content of the speaker's beliefs, in effect Davidson has shown that, as a methodological necessity for interpretation, a necessary condition for our having beliefs is that the principle of charity be employed across the board. Thus if the principle of charity is itself anti-sceptical, so much the worse for scepticism; we now seem to have in hand an argument which has brought to fruition the ambitious Kantian anti-sceptical strategy discussed above.

All is not well in this putative philosophers' paradise, however. It is suspicious that we have found no use for the additions Davidson made to the omniscient interpreter argument in its revised formulation, in particular the explicit claim that the omniscient interpreter "is omniscient about... what does and would cause a speaker to assent to any sentence in his (potentially unlimited) repertoire". As of yet, we have had no recourse to Davidson's externalism. The basic problem is that although the need to apply the principle of charity across the board shows why it is impossible for an omniscient interpreter correctly to interpret a speaker as massively mistaken, by itself this seems consistent with the possibility of everyone, mere humans and "omniscient interpreters" alike, being massively mistaken about the world.

How can this be? Recall that the posited omniscience doesn't directly extend to intensional facts; he attributes meaning and beliefs to others just as we do, viz. by interpreting their sentences. The thought behind the argument I gave was that because, first, his omniscience extends to all facts about the external world and, second, he must apply the principle of charity when interpreting our utterances, the omniscient interpreter necessarily will interpret us as in general agreement with his own beliefs, which are, by hypothesis, "objectively correct". This hypothesis rankles. Doesn't the presumption that the omniscience interpreter's beliefs are objectively correct contradict Davidson's claim that there is no interpretation-independent fact of the matter about what someone believes? It seems that a so-called "omniscient" interpreter would believe whatever we fallible interpreters would make him out to believe. Yet we have no guarantee that our standards are objectively correct; showing that they are is precisely the point of the omniscient interpreter argument. There is no reason to think that an omniscient interpreter is in any better position than we are to hold true what actually is true; when it comes to holding true only what actually is true, all interpreters are equal. It seems that the omniscient interpreter argument cannot get started since to imagine

as possible an omniscient interpreter is to beg the very question the omniscient interpreter argument is supposed to answer.

Although I argue below that this objection can be rendered harmless, I think it shows that the application of the principle of charity cannot, by itself, serve to guarantee the general veridicality of our beliefs. I have maintained throughout that externalism is essential to Davidson's anti-sceptical argument. Earlier I cited his externalist claim "we can't in general first identify beliefs and meanings and then ask what caused them" (1983: 131). Why can't we? In "Empirical Content", he explains that the reason "our basic methodology for interpreting the words of others necessarily makes it the case that most of the simplest sentences which speakers hold true *are* true" is that "the interpreter... must take into account the causal interaction between world and speaker" (1982: 322). Why must an interpreter do this? Davidson offers this explanation in "Epistemology Externalized":

If anything is systematically causing certain experiences (or verbal responses), that is what the thoughts and utterances are about. This rules out systematic error. If nothing is systematically causing the experiences, there is no content to be mistaken about. To quote myself [from Davidson, 1983: 132]: "What stands in the way of global scepticism of the senses is, in my view, the fact that we must, in the plainest and methodologically most basic cases, take the objects of belief to be the causes of that belief". Anyone who accepts perceptual externalism knows he cannot be systematically deceived about whether there are such things as cows, people, water, stars, and chewing gum. (1991: 195)

At last we see a connection between meaning and truth. However, although this doctrine seems powerfully anti-sceptical, consider the question of how we determine which cause it is that our basic thoughts are about. For instance, what cause is content-determining in the case of our thought "there's a cow"? As Davidson admits, there seem to be many potential content-determining causes, ranging from some "going back in time before all cows" to those "spatially closer to the thinker than any cow" (*ibid.*). How can perceptual externalism accommodate the fact that our thoughts about cows refer to the "natural" causes, the bovine animals? If it cannot, the doctrine must be false.

In a phrase, the problem is to understand how the "objects of thought" are to be identified. Davidson's answer, that "the identification of the objects of thought rests... on a social basis" (1991: 196), is based on an analogy

with language learning. It will be useful to quote at length from “Epistemology Externalized”:

We cannot... consider the question of the contents of mental states from the point of view of a single creature. This is perhaps best seen by thinking about how one person learns from another how to speak and think of ordinary things. Put in greatly simplified terms, a basic aspect of such learning can be described in this way: the learner is rewarded... when the learner makes sounds or otherwise responds in ways the teacher finds appropriate in situations the teacher classes together. The learner is subsequently caused to make similar sounds by situations the learner instinctively classes together... Success at the first level is achieved to the extent that the learner responds with sounds the teacher finds similar to situations the teacher finds similar. The teacher is responding to two things: the external situation and the responses of the learner. The learner is responding to two things: the external situation and the responses of the teacher. All these relations are causal. Thus the essential triangle is formed which makes communication about shared objects and events possible. But it is also this triangle that determines the content of the learner’s words and thoughts when these become complex enough to deserve the term... .[W]hat makes the particular aspect of the cause of the learner’s responses the aspect that gives them the content they have is the fact that this aspect of the cause is shared by the teacher and the learner. (1990: 197-8)

The analogy provides powerful motivation to accept the Davidsonian claim, enshrined in radical interpretation, that one can be a believer only if one is an interpreter of others. The powerful suggestion is that, as believers and communicators, we are all like the learner and the teacher: others attribute meaning and belief to us by judging certain utterances of ours as similar by correlating them with external situations they judge similar; we attribute meaning and belief to others by correlating those responses we judge similar with external situations we judge similar. The objects of thought, the “natural” causes, are determined by interpretative triangulation; in the basic cases at least, the content of our own beliefs is determined by just those external situations we, speakers and interpreters who find our utterances mutually interpretable, “naturally” find similar.

The problem of identifying the objects of thought seems to be a general problem, and one of extreme significance. Certainly when investigating occasion sentences like Quine’s observation sentences or Davidson’s “methodologically basic” sentences, all extensionalist theories of knowledge must make some appeal to observable circumstances under which sentences are

held true. But which element in the causal path ending in the speaker's mental states fixes the salient circumstances? Consider again the thought "there's a cow". Obviously, it is impossible for a speaker and interpreter to share causes "going back in time before all cows". The only plausible candidates for intermediate causes "spatially closer to the thinker than any cow" are either literally on the skin (something like "photon arrays on the surface of the retina") or just inside it (perhaps "patterns of sensory nerve stimulation"). But these things cannot be routinely observed at all; how could they be the salient shared circumstances? If what a person means can in principle be determined by another person's observations, then the problem of determining the objects of thought seems to show that externalism is unavoidable.

I have only begun to draw out the importance of the problem of determining the objects of thought; for example, what we have seen suggests reasons to avoid all "epistemic intermediary" accounts of meaning and justification. For the purposes of this essay, however, a more relevant topic is how this account of the role of externalism in radical interpretation affects the reading of the omniscient interpreter argument I gave earlier. It seems to me that the interpretation was largely correct, but it missed the main point of the argument. Recall this crucial claim:

The omniscient interpreter, using the same method as the fallible interpreter, finds the fallible speaker largely consistent and correct. By his own standards, of course, but since these are objectively correct, the fallible speaker is seen to be largely correct and consistent by objective standards. (1983: 131)

The objection I raised was that there can be no great difference between the standards of omniscient and "merely fallible" interpreters. I took this to show that assuming the possibility of an omniscient interpreter with objectively correct standards begged the question. In the light of the doctrine of perceptual externalism, however, it becomes clear that the standards employed by any interpreter, human or omniscient, are of necessity objectively correct. On this reading, the omniscient interpreter argument breaks no new anti-sceptical ground; perhaps it is best seen as expressively drawing out some of the consequences of the externalistic claims that (at least in Davidson's more recent formulation) precede and follow it.

### 5. Some objections and replies

At this point, one can predict a sceptic's complaint that "if *that* is Davidson's doctrine, then by accepting perceptual externalism Davidson begs the question plain and simple; he just hasn't taken scepticism seriously." On the contrary, in this case it would be the sceptic who fails to take externalism seriously. Given the extensionalist stance that "what a fully informed interpreter could learn about what a speaker means is all there is to learn" and "... the same goes for what the speaker believes" (1983: 129), all conditions necessary for interpretation are also necessary conditions for having belief. I argued earlier that although perhaps "global nihilists" can freely question them, global sceptics must respect all the necessary conditions for belief; we saw that the sceptical question itself makes sense only if these conditions are taken for granted. Finally, we have seen why externalism is essential for radical interpretation. If externalism weren't true, then, we would have no way of determining the objects of thought in the most simple, "methodologically basic" cases. Davidson has argued not that externalism presupposes the falsity of scepticism, although this is true. Rather, his argument is that the sceptical position itself must, inconsistently, presuppose externalism.

Having said this in favour of Davidson's anti-sceptical strategy, I would like to raise a potential objection based on the thought that externalism is incompatible with the phenomena of first person authority. In "Epistemology Externalized", Davidson characterises externalism as the doctrine "which holds that the contents of a person's propositional attitudes are partly determined by factors outside the mind *of which the person may be ignorant*" (1990: 193; my emphasis). How is it possible to know the content of one's own thoughts while being ignorant of their causal antecedents? A worry is that if it is the actual causal antecedents which determine meaning, my thought "there's a rat" would have no fixed meaning: usually it will be a thought about rats, but sometimes, depending on what actually causes me to hold it true, it is about rat-shaped shadows or giant cockroaches. The problem is that, without presupposing just the knowledge that the global sceptic denies we possess, it seems impossible to know what our thoughts are about at any given time. Moreover, if the content of our thoughts varies along with their actual causes, our simplest thoughts about the world can never be wrong: on the one hand, if there is no external cause, they have no content;

on the other hand, if they have content, our thoughts cannot but be about whatever external circumstances in fact cause them.

Davidson has an elegant argument that externalism is perfectly compatible with first person authority:

An interpreter must discover, or correctly assume on the basis of indirect evidence, what the external factors are that determine the content of another's thought; but since these factors determine both the contents of the thought and the contents of the thought one believes one has, there is no room for error about the contents of one's own thoughts of the sort that can arise with respect to the thoughts of others. (1990: 194)

Since the same external factors determine "both the contents of the thought and the contents of the thought one believes one has", the doctrine of externalism guarantees that the content of our thoughts about our beliefs, e.g., "I believe that I saw a rat," necessarily must match the beliefs they report, e.g., "there's a rat". For this reason, no independent knowledge of those external factors is required to know the content of one's own thoughts. But do the contents of those thoughts really vary in the absurd way described above, viz. sometimes being determined by one thing, sometimes by another, and sometimes by nothing at all (in which case the first-level and the corresponding second-level beliefs have no content)? Not at all. A deeper problem with this objection is that it fails to comprehend the importance of the language learning analogy in Davidson's triangulation argument.

Davidson's externalism is not the doctrine that what our thoughts mean is determined by whatever in our environment happens to prompt them *on each occasion*. For Davidson, language learning is not just a handy metaphor; in the basic cases, the meaning of our thoughts is fixed by such a process. It is those causes, literally shared by a teacher and a language-learner, that determine what our thoughts mean *from that time forth*. Error thus has a simple explanation: if, after we have acquired concepts and beliefs about mice, we subsequently think "there's a mouse" in the absence of whatever originally fixed the relevant concepts, then what we say is false.

This suggests a way to handle brain-in-vat objections. In the standard case, the allegedly sceptical claim is that I might be an envatted brain connected to a stimulation computer that stimulates my sensory nerves in just the way they would be if I were embodied. One question to consider is whether it makes sense to attribute beliefs to envatted brains. We have seen

that it is a central Davidsonian claim that meaning is possible only if a radical interpreter could “discover... what the external factors are that determine the content of another’s thoughts” (1990: 193). We saw that the objects of thought can be identified only if there is the possibility of a rather complex triangular interaction between speaker, interpreter, and the world. Can an envatted brain interact in the way required for it to be correctly interpreted as having beliefs? If not, it follows that, since we do have thoughts, we cannot be envatted brains. However, it seems likely that it would be possible to come up with some variation on the brain-in-vat story that meets the intersubjective conditions on thought. Let’s assume that something along the following lines will suffice: a brain has been envatted as in the standard case, with the addition, first, of some (amplification and other) equipment through which it can communicate with an interpreter, and, second, of some sort of computer-fed stimuli which are themselves caused by the interpreter’s inquiries.

I think that Davidson’s externalism provides a quick way to handle this case. Since they are supervenient on (among other things) their actual causes, mental states are not reducible to “stimulus-identical” brain states. As the radical interpreter would attest, the envatted brain’s utterances are systematically correlated with states of the stimulation computer; it follows that although in the envatted and embodied cases the brain would (by hypothesis) be in the same physical brain states, its mental states would be different. The point is that, in envatted circumstances, the brain’s report that “I have thoughts about a world external to me and which includes such-and-such enduring, interacting objects” does not mean what it would in the embodied case; in the envatted case such thoughts are true statements about the envatted brain’s environment. Since a brain’s “thoughts about the world” are about their interpretable causes, the possibility that I am an interpretable brain in a vat carries with it no sceptical consequences; the contents of my thoughts about the world are determined by my actual environment.

Now consider a slightly different brain-in-vat story. While she was asleep, Sally’s embodied brain was envatted; when Sally regains consciousness she utters “there’s a mouse”. But Sally does not see a mouse; the stimulation computer has simply caused her to have the same brain state she would have had if embodied and seeing a mouse. In this case, it seems that Sally is in error: after having acquired concepts and beliefs about mice in the normal way, she subsequently thinks “there’s a mouse” in the absence of

what originally fixed her concepts. Although in this case Sally still knows many things about the world (that there are mice, that they eat cheese and are eaten by cats, etc.), she seems doomed to acquire a huge quantity of false beliefs through her connection with the stimulation computer. Even though Sally is better off epistemically than the global sceptic would like to admit, her position doesn't seem that much better. However, reflect on how Sally's utterances would come to be interpreted. After observing Sally's envatted brain, a radical interpreter will come correctly to interpret her thoughts as being about states of the stimulation computer. If immediately upon regaining consciousness her thought "cats eat mice" was a true thought about cats and mice, presently it will come to be interpreted as a true thought whose content is determined by her envatted environment. Thus Davidson endorses Tyler Burge's argument that "such a brain cannot (for long) be radically deceived about its environment" (Davidson, 1990: 195).

Consider, finally, a very different brain-in-vat story. Consider a hapless brain which undergoes an endless changes of environments: first it is embodied, then it is envatted, then it comes under the spell of Cartesian demonology, then it is envatted again, etc. We saw that Sally came to be briefly radically deceived by the computer's stimuli. Suppose that in this case the brain changes from one environment to the next so rapidly that its concepts never have time to "harmonise" with its environment. Wouldn't this unfortunate brain always be massively deceived about its environment? I think, rather, that a brain in this state would quickly become uninterpretable. After all, there never was a time during which Sally continued "incorrectly" to use her embodied concepts; at each moment, Sally means whatever it is a radical interpreter would correctly make her out as meaning. Immediately after becoming disembodied, Sally's brain would be correctly interpreted as thinking in English about the very things she had known while embodied; in time, however, as more and more of her mental states came to be caused by the computer, the interpreter (employing the principle of charity) would correctly interpret her as having thoughts about her envatted environment. Similarly, when it acquired a significant number of mental states caused by non-standard environments, the poor brain subjected to constant changes of environment could no longer be correctly interpreted as thinking in English. But whereas the envatted Sally had a stable causal environment about which she could be interpreted as thinking, lacking this the other brain would simply become uninterpretable. As we saw, according to Davidson's

doctrine of externalism “if nothing is systematically causing the experiences, there is no content to be mistaken about” (1990: 5).

### **6. A modest result**

I started this paper fully expecting to find Davidson’s response to global scepticism fundamentally flawed. With my critique of Kant’s anti-sceptical argument in hand, I anticipated quickly finding a question-begging move. Absent this, it seemed likely that his strange doctrine of externalism would be susceptible to new sceptical challenges; difficulties with our knowledge of our own thoughts seemed to point towards a new “scepticism of mental content” and sceptical puzzles involving baroque brain-in-vat stories loomed large. I now believe that Davidson’s doctrine of externalism is well-equipped to frustrate these objections. Combined with the confirmation of my suspicions about the affinities between Davidson’s and Kant’s approaches to scepticism, this was a pleasing result to me; my earlier look at Kant had convinced me of the great potential of his ambitious anti-sceptical project.

Davidson advertises his work on meaning and knowledge as contributing to nothing less than “a sea change in contemporary philosophical thought” (1989: 159). The production of a good anti-sceptical argument truly would be a revolutionary achievement. Although reflecting on the problem of determining the objects of thought appears to go a long way towards vindicating his externalism, I have left untouched too many aspects of Davidson’s position — no doubt in the process running roughshod over many important points and missing many potential objections — to be in a position to hazard a final assessment of his answer to the sceptical question. Nonetheless, it is an important result just to have shown that Davidson’s externalism qualifies as a genuinely anti-sceptical position, i.e. one that might well be the basis of a decisive, knockdown argument against global scepticism. This result rewards my interest in Davidson and Kant and provides a glimpse of exciting new vistas afforded by further inquiry into “epistemology externalized”.

Andrew Carpenter  
Dept. of Philosophy  
University of California at Berkeley  
1290 Laurel Lane, Westminister, MD 21158, USA  
phlos-ad@socrates.berkeley.edu

*References*

- Brueckner, A. (1986) "Charity and skepticism" in *Pacific Philosophical Quarterly* 67, pp. 264-268.
- Burge, T. (1986) "Cartesian error and the objectivity of perception" in P. Pettit and J. McDowell (eds.), *Subject, Thought, and Context*. Oxford: Clarendon Press, pp. 117-136.
- Davidson, D. (1967) "Truth and meaning" in D. Davidson, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, pp. 17-36.
- Davidson, D. (1973) "Radical interpretation" in D. Davidson, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, pp. 125-140.
- Davidson, D. (1974) "Belief and the basis of meaning" in D. Davidson, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, pp. 141-154.
- Davidson, D. (1974b) "On the very idea of a conceptual scheme" in D. Davidson, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, pp. 183-198.
- Davidson, D. (1975) "Thought and talk" in D. Davidson, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, pp. 155-170.
- Davidson, D. (1976) "Reply to Foster" in D. Davidson, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, pp. 171-180.
- Davidson, D. (1977) "The method of truth in metaphysics" in D. Davidson, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, pp. 199-214.
- Davidson, D. (1977b) "Reality without reference" in D. Davidson, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, pp. 215-226.
- Davidson, D. (1979) "The inscrutability of reference" in D. Davidson, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press, pp. 227-242.
- Davidson, D. (1982) "Empirical content" in E. LePore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell, pp. 320-332.
- Davidson, D. (1982b) "Rational animals" in *Dialectica* 36, 317-327.
- Davidson, D. (1983) "A coherence theory of truth and knowledge" in A. Malachowski (ed.), *Reading Rorty*. Oxford: Basil Blackwell, pp. 120-134.
- Davidson, D. (1984) "First person authority" in *Dialectica* 38, 101-111.
- Davidson, D. (1987) "Afterthoughts to 'A Coherence Theory of Truth and Knowledge'" in A. Malachowski (ed.), *Reading Rorty*. Oxford: Basil Blackwell, pp. 134-137.

- Davidson, D. (1987b) "Knowing one's own mind" in *Proceedings and Addresses of the American Philosophical Association*, pp. 441-458.
- Davidson, D. (1989) "The myth of the subjective" in M. Krausz (ed.), *Relativism: Interpretation and Confrontation*. South Bend, Illinois: Notre Dame Press, pp. 159-172.
- Davidson, D. (1989b) "The conditions of thought" in J. Brandl and W. Gombocz (eds.), *The Mind of Donald Davidson*. Amsterdam: Rodopi, pp. 193-200.
- Davidson, D. (1990) "Meaning, truth, and evidence" in R. Barrett and R. Gibson (eds.), *Perspectives on Quine*. Oxford: Basil Blackwell, pp. 40-64.
- Davidson, D. (1991) "Epistemology externalized" in *Dialectica* 45, 191-202.
- Davidson, D. (1991b) "Three varieties of knowledge" in A.P. Griffiths (ed.), *A.J. Ayer: Memorial Essays*. Cambridge: Cambridge University Press, pp. 153-166.
- Davidson, D. (1992) "The second person" in P. French and T. Uehling (eds.), *Midwest Studies in Philosophy 17: The Wittgenstein Legacy*. Notre Dame: Notre Dame University Press, pp. 255-267.
- Davidson, D. (1995) "Pursuit of the concept of truth" in P. Leonardi (ed.), *On Quine: New Essays*. New York: Cambridge University Press, pp. 203-220.
- Klein, P. (1986) "Charity and global skepticism" in E. LePore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell, pp. 369-386.
- McGinn, C. (1986) "Radical interpretation and epistemology" in E. LePore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell, pp. 356-368.
- Rovane, C. (1986) "The metaphysics of interpretation" in E. LePore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell, pp. 417-429.
- Stroud, B. (1984) "Naturalized epistemology" in B. Stroud, *The Significance of Philosophical Scepticism*. Oxford: Oxford University Press, pp. 209-254.
- Vermazen, B. (1983) "The intelligibility of massive error" in *Philosophical Quarterly* 33, 69-74.
- Williams, M. (1988) "Skepticism and charity" in *Ratio (New Series)* 1, 176-194.