

SELF-KNOWLEDGE AND EXTERNALISM

BILL BREWER

I want to discuss the possibility of combining a so-called ‘externalist’ theory of empirical content, on which the contents of a person’s beliefs are determined in part by the nature of his extra-bodily environment, with a plausible account of self-knowledge, in particular, of a person’s knowledge of the contents of his own beliefs. A difficulty for this combination is thought to be that it provides a wholly non-empirical source of knowledge about the mind-independent physical world which is intuitively intolerable.¹ The inference which is held to create this difficulty can be put like this.

- (E1) I believe that p.**
- (E2) If x believes that p, then x ’s environment is thus-and-so.**
- ∴ (E3) My environment is thus-and-so.**

An example of this might be the following.

- (w1) I believe that water is wet.**
- (w2) If x believes that water is wet, then x ’s environment contains (or did contain) water.**
- ∴ (w3) My environment contains (or did contain) water.**

¹The particular formulation of the issue which I address here is due to Martin Davies (1997). Indeed, my thoughts in this area were initially stimulated by the opportunity to give a reply to an earlier version of his paper at a meeting of the European Society for Philosophy and Psychology in Barcelona during the summer of 1996. Paul Boghossian also presses the issue as a challenge to content externalism of the kind defined in the text below (1989, 1997).

Note that the truth of the consequent of (w2) requires that x's environment contains (or did contain) water, rather than any kind of 'twin-water', which is like water in all superficial respects but happens to have a different chemical composition - that is, it requires that x's environment contains (or did contain) this stuff (H₂O) as opposed to that stuff (XYZ) (Putnam, 1975). This is precisely the force of the relevant form of content externalism. So the conclusion states that my environment contains water, as opposed to twin-water. This is a contingent matter of empirical fact, though. Hence the prospect of my knowing it without any kind of empirical investigation, certainly raises a prima facie problem.

I

The proponent of this line of objection to combining an adequate account of self-knowledge with content externalism argues as follows. First, any adequate account of a person's knowledge of the contents of his own beliefs entails that his knowledge of instances of (E1) is non-empirical: neither its acquisition nor its status as knowledge necessarily involves any specific empirical investigation. Second, content externalism entails that instances of (E2) can be derived from non-empirical philosophical reflection, on the necessary conditions upon determinate empirical belief possession. Third, therefore, the truth of content externalism - in the presence of an adequate account of self-knowledge - enables a person knowledgeably to derive instances of (E3), on the basis of the argument above, without any empirical investigation whatsoever. Fourth, such non-empirical knowledge of empirical facts is intuitively intolerable. Therefore, fifth, content externalism is incompatible with any adequate account of a person's knowledge of the contents of his own beliefs. The reaction which is implicitly recommended by advocates of this argument, of course, is to reject content externalism. Like many others, though (e.g. Putnam, 1975; Burge, 1979, 1986; Evans, 1982; Pettit and

McDowell, 1986; McGinn, 1989; Davies, 1991, 1992; Peacocke, 1993), I am very sympathetic to some form of externalism. Indeed, I think that world-dependent perceptual demonstrative contents are crucial, both to our very capacity to think about particular things in the world around us, and to our acquisition of knowledge about them (1999). So I need some alternative to the recommended reaction of rejecting such externalism.

Before considering what this should be, it is worth pointing out that some externalists have tried to exploit the intuitively intolerable result directly to their own advantage in presenting anti-sceptical transcendental arguments. “There is no such thing as a reductio ad absurdum”, they say, “just the consequences of my view; and this one is very useful thank you”. For example, Burge offers the following in this spirit.

(B1) I think that there are physical entities.

(B2) If x thinks that there are physical entities, then x bears causal-perceptual relations to physical entities.

∴ (B3) I bear causal-perceptual relations to physical entities.

I shall return to my assessment of this bullish attempt to turn the tables on critics of externalism when I have given my own account of how to respond to the apparent threat posed by arguments along these lines. Cutting to the chase: Burge’s attempt fails.

What, then, are the externalist’s options here? **The one which I shall eventually propose and defend** is to reject the very first move of the objection set out above. I deny that a person’s knowledge of the contents of his own beliefs is entirely non-empirical in any way in which it really would follow that he could thereby acquire intolerably non-empirical knowledge of the contingent facts about his own environment. There is, in my view, an essential empirical component in self-knowledge. For determinate concept

possession is, in a sense which I shall make precise, an epistemic skill. Before developing this claim, I want briefly to survey the alternatives.

First, it might be possible to argue that the content externalist is not committed to the non-empirical knowability of specific instances of (E2). Although he is committed by definition to the conceptual necessity of the claim that concepts of certain types are externally individuated, in a way which entails the possibility of establishing that claim by non-empirical reflection, he may nevertheless deny that the question of which particular concepts instantiate these types can be settled without empirical investigation.²

Second, it might also be possible to deny the third move above, according to which the truth of content externalism - in the presence of an adequate account of self-knowledge - enables a person knowledgeably to derive instances of (E3) without any empirical investigation. This is Martin Davies' response (1997), which I shall also consider in some detail shortly.

There are two further responses, which I shall simply mention and set aside without argument at this stage, as extremely unpromising. The first of these (3) would be to insist, in extreme rationalist spirit, that there is no difficulty at all in the idea of wholly non-empirical knowledge of empirical matters of fact. The final status of the contrary intuition, against this possibility of natural science by non-empirical reflection, is bound up, of course, with the fate of anti-sceptical transcendental arguments like Burge's, above, to which I shall return. The second unpromising response (4) would be to claim that a person's self-ascriptions of beliefs are not, contrary to appearances, genuinely truth-evaluable statements, but rather non-truth-evaluable avowals of some kind (Wittgenstein, 1958, pp. 190-192; 1980, §§ 470-504; Malcolm, 1991). Thus, they are incapable of constituting the premises of an argument, as they are supposed to do in (E1),

²Christopher Peacocke urged me to take account of this possibility in correspondence.

the first premise of the argument which is in turn supposed to cause trouble for content externalism. Again, this move seems desperate at this point. Whatever their logical and epistemological peculiarities, self-ascriptions of beliefs surely are statements about a person, the person making them, to the effect that he or she is in a certain condition, namely that of believing that *p*, say.

II

Let me return, then, to the two more plausible alternatives identified above to my own preferred strategy for establishing the compatibility of self-knowledge and content externalism in the face of the present line of objection. The first of these denies that specific instances of (E2) are knowable non-empirically, whilst granting that content externalism itself is, by arguing that it is a matter of empirical investigation to which actual concepts this externalism applies. For the standard externalist examples are natural kind concepts. Yet knowledge that any given concept is a genuine natural kind concept requires knowledgeably ruling out the following two ways in which it might fail to be so, which is surely an empirical matter.

First, what appears to be reference to a natural kind may fail to be so because the characteristic theoretical role associated with the putative kind is not in fact played by anything at all. ‘Phlogiston’ provides a familiar actual case of this type. There just is no kind of stuff which is released into the surrounding medium quickly by burning and slowly by rusting, and so on. These processes involve instead the absorption of oxygen from the atmosphere. So there is no such thing as phlogiston. Apparent reference to a natural kind by the use of that term is merely illusory, and therefore certainly not susceptible to externalist treatment. Call this the possibility of emptiness with respect to a given putative natural kind term.

Second, the range of items to which a term is correctly applied may turn out to be quite heterogeneous at the relevant level of theoretical description involved in the discrimination and unification of natural kinds. 'Jade' provides a familiar actual example of this phenomenon; for it is correctly used of two quite different substances: jadeite, which is a silicate of sodium and aluminium; and nephrite, which is a silicate of lime and magnesia. Cases of this kind could clearly turn out to be far more wildly heterogeneous, sufficiently so to make quite implausible any attempt to think in terms of a single unified natural kind with a number of sub-varieties, as it may be correct to think that some water is H₂O and some is D₂O, although both are sub-varieties of the single natural kind water, and not of the natural kind twin-water (XYZ). In the more extreme cases, again, what appears to be a natural kind term turns out, on empirical investigation, not to be so. Call this the possibility of heterogeneity.

Thus, it may be claimed, knowledge of any particular instance of (E2) depends upon knowledge that the relevant component concept of the content that p is a genuine natural kind concept, which in turn depends upon the empirical knowledge required to rule out the possibilities of emptiness and heterogeneity with respect to the term expressing that concept. So the argument from (E1) and (E2) to (E3) does not, after all, threaten to provide an untenable, because wholly non-empirical, source of any specific piece of empirical knowledge.

I think that the content externalist should not be overly impressed by this apparent resolution of the tension between his position and the possibility of a plausible account of self-knowledge. The way in which parallel issues arise in connection with perceptual demonstrative reference to particular objects helps to bring out why. Consider, for example, the following instance of the argument (E1), (E2) |- (E3).

- (d1) I believe that that ball will go into the pocket.**
- (d2) If x believes that that ball will go into the pocket, then x's environment contains that ball.**
- ∴ (d3) My environment contains that ball.**

Of course, there is no threat that such an argument should constitute an untenably non-empirical source of empirical knowledge - that a particular ball exists in his environment - if the subject's attempted demonstrative reference fails. And there are at least the following two ways in which this might come about. First, there may be no ball at all, or anything else for that matter, at the place in his environment where the subject takes there to be one. That is to say, he may be subject to some kind of hallucination. Second, his thinking may fail to be responsive to the behaviour of a single ball, as he fails, for example, to keep track of the movement of a single red ball after a very strong break in a game of snooker. In the absence of any such defeating abnormality, though, the argument is perfectly sound; and neither his successful demonstrative reference to the particular ball in question, nor his knowledge of the particular instance of (E2) - that is, his knowledge of (d2) above - depends upon his having carried out any prior, or independent, empirical investigation to rule out all the possible sources of hallucination, unnoticed substitution of one ball by another, failure of attentional tracking, and so on.

Similarly, I contend, in the case of reference to natural kinds. According to the most natural development of the externalist position, understanding of natural kind terms is, in the most basic cases, acquired by some kind of demonstrative identification of the kind in question, as 'that stuff' - 'that liquid', say - on the basis of perception of its instances and some grasp their distinguishing features and characteristic behaviour. The possibilities of emptiness and heterogeneity with respect to putative natural kind terms

arise as follows. In the first case, the appearance of an underlying kind is entirely hallucinatory. In the second case, the attempted demonstrative identification fails sufficiently to keep track of any single kind. Either way, the relevant instance of the argument (E1), (E2) |- (E3) is quite harmless to the content externalist: it's second premise is simply false. Nevertheless, in the normal case, when the demonstrative identification crucial to the subject's understanding of the term in question makes successful reference to a genuine natural kind, neither this, nor his knowledge of the corresponding specific instance of (E2), depends upon his having carried out any prior, or independent, empirical investigation to rule out all the possible sources of error due to emptiness or heterogeneity. Thus, the relevant instance of (E1), (E2) |- (E3) is back in contention as a threat to content externalism, appearing, as it does, to provide an untenably non-empirical source of empirical knowledge. In any case, the present attempt to avoid this difficulty for combining content externalism with an adequate account of self-knowledge is unsuccessful.

The second prima facie promising alternative to my own account of how the two are to be combined is to appeal to Davies' (1997) claim that, regardless of the validity of the argument from (E1) and (E2) to (E3), of a person's knowledge of instances of its two premises, and, indeed, of his knowledge of the validity of the relevant form of argument itself, knowledge of instances of its conclusion still cannot be acquired by these means. For this would be in contravention of the following Limitation Principle.

(LP) Epistemic warrant cannot be transferred from A to B, even given an a priori known entailment from A to B, if the truth of proposition B is a precondition of the knower being able to believe the proposition A.

The only motivation which he offers for this principle is that it is supposed to save him from precisely the present difficulty with reconciling self-knowledge and

externalism; and it looks a little bit like the following principle which he reads into Wittgenstein's On Certainty (1975).

(WP) Epistemic warrant cannot be transferred from A to B, even given an a priori known entailment from A to B, if the truth of proposition B is a precondition of the warrant for A counting as a warrant.

I think that neither is what Wittgenstein has in mind. More importantly, neither is remotely plausible in my view: (LP) even less so than (WP). A whole class of serious counterexamples to (LP) come immediately to mind. Consider the following inference.

(M1) r_1 and r_2 are real numbers

(M2) The product of any two real numbers is a real number

\therefore (M3) $r_1 r_2$ is a real number

Its conclusion is a necessary truth, which is therefore, presumably, a precondition of anything possible. Hence it is certainly a precondition of a person's believing the premises (M1) and (M2). So it follows from (LP) that epistemic warrant cannot be transferred from (M1) and (M2) to (M3). Generalizing the case, (LP) has the consequence that inferential knowledge of any necessary truth is impossible, which certainly places logic, mathematics, and, on many conceptions, philosophy, in a very poor position epistemologically speaking. Something must have gone badly wrong with (LP). So Davies' escape route for the externalist is closed.

III

The key to my own strategy for reconciling self-knowledge with content externalism lies in the idea that true content externalist requirements are a consequence of the following version of Russell's Principle of Acquaintance (1917, p. 159), which I shall call (A).

(A) A person's capacity to make determinate reference to certain objects and kinds in belief depends upon his having demonstratively-based knowledge about them.

This is what makes it the case that his possession of such beliefs depends upon his being in an environment actually containing the objects and kinds in question. What does this acquaintance condition amount to, though? Well, a person has demonstratively-based knowledge about a given object or kind just if, either, he has knowledge expressible using demonstrative reference to that object or kind, or, he had such knowledge, some of which is retained, perhaps linguistically categorized and no longer demonstratively expressible, in memory.

Establishing this principle (A) obviously requires extended argument, of which I only offer a sketch here.³ The crucial claim is that externalist relations are necessarily reason-giving relations, constituting a source of demonstratively expressible knowledge; where by this I mean reason-giving in the epistemological internalist's sense: reason-giving from the subject's point of view, rather than merely from the perspective of some external theorist. Let me explain what this means and why I endorse it.

The externalist holds that certain empirical beliefs have their contents determined in part by the subject's relations with particular things in his environment, the causal

³The argument is developed in detail, given extended illustration, and defended against putative counterexamples in ch. 3 of my forthcoming book (1999).

relations, for example, in which he stands to such things - to that object, say, or (instances of) that natural kind - when he is perceiving them, being informed about them by others, remembering them, and so on. Call these his causal-perceptual relations with the external worldly things in question. Now, suppose that these content-determining causal-perceptual relations are not reason-giving relations; and consider a person, (our old friend) S, who believes that p (as usual), where this is supposed to be an empirical belief with externalistically determined content. Since his causal-perceptual relations with the things around him play an essential role in the determination of the contents of his empirical beliefs, on the externalist's account, there is a range of alternative such beliefs - beliefs which he might have had instead - whose difference in content with his actual belief that p would have been due entirely to his standing in the relevant content-determining causal-perceptual relations with different mind-independent things. Suppose that the belief that q is one of these.

So, the situation is this. S actually believes that p, because his actual environment determines this, as opposed to q, as the empirical content of his belief, through its causal-perceptual impact upon him. He does not believe the distinct content that q. Had his environment been appropriately different, though, his position would have been precisely the reverse: he would have believed that q, and not believed that p. Yet, by hypothesis, the relevant content-determining causal-perceptual relations in which he stands to the actual things in the world around him are not reason-giving relations. So his standing in them is quite neutral, from his point of view, on which such things it is to which he is so related. For if they were not neutral in this way, then they would be reason-giving in the relevant sense. Thus, his standing in these relations is quite neutral, from his point of view, on whether it turns out that he is believing that p or that q. Yet they are the only difference between the actual case, in which he believes that p, and the counterfactual case, in which he believes that q - for this is how q was introduced in relation to p. Thus,

there could not possibly be, on this account, any difference at all, from the subject's point of view, between his believing that *p* and his believing that *q*.

Which of the two beliefs he actually has is due entirely to which environmental entities it happens to be in which he stands in the relevant causal-perceptual relations; yet these relations are quite neutral, from his point of view, on which entities these turn out to be. Any supposed difference between believing that *p* and believing that *q* is therefore absolutely nothing to him. So he does not really understand them as alternatives.

Believing that *p* and believing that *q* are identical for him. Hence the supposedly content-determining role of *S*'s environment is empty. For there is nothing more, or less, to the content of a belief than the way the subject takes the world to be. Thus, if the proposed causal-perceptual relations in which a person stands to certain mind-independent things are not reason-giving relations, then they contribute nothing to the determination of specific truth-conditions for his empirical beliefs. In other words, the content-determining relations between a person and certain things in the world around him which are posited by the content externalist must be reason-giving relations.

The form of this argument is that of what Peacocke calls 'the switching tactic' (1988, pp. 475 ff).⁴ A more familiar historical paradigm is provided by Strawson's (1959, ch. 3; 1966, pt. III, sect. II, esp. pp. 168 ff; and 1974) reading of Kant's (1929, A341/B399 ff) objection to Descartes' (1986, pp. 107 ff) substance dualism. According to this Kantian argument, substance dualism entails the coherence of a distinction between qualitative and numerical identity for immaterial minds; yet the dualist's own conception of such things, as exhaustively characterized by what is infallibly given to their own subjective point of view, denies her the resources to give any genuine content

⁴So-called, I presume, because the tactic is to object to a theory on the grounds that it is in principle incapable of giving any significance to switches which are by that theory's own lights crucial.

to the idea of two qualitatively identical but numerically distinct minds. So, the substance dualist depends upon a distinction - between qualitative and numerical identity for immaterial minds - which she is, by her own lights, incapable of making. Thus, the position is internally inconsistent.

Similarly, here, we are to consider a theorist who insists that the non-reason-giving causal-perceptual relations in which a person's stands to certain objects or kinds in his worldly environment are essential to the determination of specific contents for his empirical beliefs. Such a theorist is therefore committed to the existence of pairs of beliefs with genuinely distinct contents, the distinction between which is entirely due to their subjects' standing in the relevant causal-perceptual relations with different possible objects or kinds. Given her own conception of the nature of these content-determining relations, though, as non-reason-giving relations, this entails an overall conception of belief content which countenances the following situation. A person actually believes that p (and does not believe that q). He might instead have believed that q (and not believed that p). Yet everything involved in determining the actual case as one of his believing that p, not that q, is entirely neutral, from the subject's point of view, between this situation and the reverse, in which he believes that q, not that p. In other words, non-reason-giving content-determining relations of this kind induce distinctions in belief content which are more discriminating than anything which could possibly enter into the subject's own understanding of his beliefs. The content of a belief is precisely, no more and no less than, the way the subject understands things to be. Yet the theorist under consideration here is committed incoherently to making distinctions in belief content which she is incapable by her own lights of making in these terms.

There is a line of thought in the recent externalist literature which is supposed to address just this kind of worry (see Burge, 1996; and Peacocke 1996). It suggests that a

person's own grasp of his externally determined belief contents consists in his ability authoritatively to self-ascribe precisely those contents, whichever they turn out to be. Suppose that S's actual belief that p and counterfactual belief that q would each be expressed by the sentence 'K is F'. Their difference in content resides in the fact that the kind term 'K' has a different semantic value in each case, where the semantic value of such a term is the natural kind in the world whose association with that term determines its contribution to the truth-conditions of sentences expressed using it. In the actual case, 'K' has semantic value k_1 ; in the counterfactual case, its semantic value is k_2 . Now, the current line of thought is that, since the same term is used in the subject's self-ascription as is used in expressing the first order belief which he thereby self-ascribes, this self-ascription is bound to be correct in either case. So his capacity authoritatively to make it constitutes his understanding of his actual belief that p, in a way which distinguishes this from the counterfactual belief that q. In self-ascribing as follows: 'I believe that K is F', in the actual case, in which the semantic value of 'K' is k_1 , he correctly self-ascribes his belief that p and not the belief that q, although he equally correctly self-ascribes the distinct belief that q, in the counterfactual case in which 'K' has semantic value k_2 . So the difference between the cases in which he believes that p and believes that q does, after all, show up from his point of view. He distinguishes them by being able correctly to make different self-ascriptions in each.

I find this line of thought quite unconvincing. Unless he already knows which kind K actually is - that the semantic value of 'K' is k_1 as opposed to k_2 - he cannot hope to inform himself of this just by reusing the word in an attempt to tell himself which it is. This so-called 'conceptual redeployment' routine is effective only if the subject can distinguish the two self-ascriptions which he is supposed to make, in the actual and counterfactual cases. Yet this presupposes his prior grasp of the distinction between believing that p and believing that q. So it cannot be brought in after the event in an

attempt to account for this understanding of the distinct first order beliefs. The subject's insensitivity to the semantic value of 'K' - that this is k_1 as opposed to k_2 - at the first order level is simply recycled in his second order self-ascription, which he therefore equally fails to understand. Although the sentence he uses to self-ascribe his belief will, on this view, be true in both the actual and counterfactual cases, the subject himself is left hopelessly ignorant of which truth this turns out to be in either case.

So, the Switching Argument stands. If a person's causal-perceptual relations with mind-independent objects and kinds are to contribute essentially to the determination of the empirical contents of his beliefs, then his standing in these relations must provide him with reasons for such beliefs: beliefs, that is to say, about just that object, or that kind. These content-determining causal-perceptual relations therefore constitute a source of demonstratively expressible knowledge about such objects and kinds: the very objects and kinds which are the semantic values of the relevant externalist concepts. Hence the source of externalist conditions upon determinate empirical belief possession is indeed my version of Russell's Principle of Acquaintance, (A) above. The requirement upon possession of empirical beliefs with certain contents that the subject should be in an environment actually containing certain objects or kinds is derived from the fact that his possession of such beliefs depends upon his actually standing in certain basic reason-giving relations with such things, which serve as a source of demonstratively-based knowledge about them.

I think that considerations from a slightly different area lend additional support to (A). Recall, first, Evans' comments about a person's use of the proper name 'Louis' in the course of his discussion of the causal theory of names (1985a).

A group of people are having a conversation in a pub, about a certain Louis of whom S has never heard before. S becomes interested and asks: 'What did Louis

do then?’ There seems to be no question but that S denotes a particular man and asks about him. Or on some subsequent occasion S may use the name to offer a new thought to one of the participants: ‘Louis was quite right to do that.’ Again he clearly denotes whoever was the subject of the conversation in the pub. This is difficult to reconcile with the Description Theory [on which there is supposed to be associated with each name as used by a group of speakers who believe and intend that they are using the name with the same denotation a description or set of descriptions cullable from their beliefs which an item has to satisfy to be the bearer of that name] since the scraps of information which he picked up during the conversation might involve some distortion and fit someone else much better. Of course he has the description ‘the man they were talking about’ but the theory has no explanation for the impossibility of its being outweighed.

The Causal Theory [on which it is sufficient for someone to denote x on a particular occasion with a name, that this use of the name on that occasion be a causal consequence of his exposure to other speakers using the expression to denote x] can secure the right answer in such a case but I think deeper reflection will reveal that it too involves a refusal to recognize the [Wittgensteinian] insight about contextual definition [that for an item to be the object of some psychological state of yours may be simply for you to be placed in a context which relates you to that thing] For the theory has the following consequence: that at any future time, no matter how remote or forgotten the conversation, no matter how alien the subject matter and confused the speaker, S will denote one particular Frenchman - Perhaps Louis XIII - so long as there is a causal connection between his use at that time and the long distant conversation. (pp. 6-7)

Evans has two important points here: first, that possession of a uniquely identifying definite description is unnecessary for successful singular reference; second, that the mere existence of a causal chain of ‘reference-preserving’ links back to the object in question, as these are conceived, for example, by Kripke (1980) and other proponents of the ‘causal theory of reference’, is insufficient. Our intuitions about the ‘Louis’ case surely confirm both of these points. My hypothesis is that these intuitions are organized

and controlled precisely by the existence of epistemic constraints upon successful reference in thought. What makes S's context in the pub conversation sufficient for him to denote Louis XIII in this way, is that he is there, at that time and in the context of that conversation, in possession of demonstratively-based knowledge about that man. His grasp of what is being said by those around him, and his understanding engagement in the discussion generally, provide him with knowledge expressible using demonstrative reference to the person in question: 'that guy (Louis) had a hard time', say.⁵ Equally, I contend, what denies his later uses of any such significance, in the circumstances which Evans describes, is that he then no longer retains anything of this knowledge in memory.

The point here can be generalized. One source of illumination for the relation of reference - both to particular objects and to natural kinds - which holds between certain referring expressions, as they are used in a given linguistic community, and the things to which they refer, lies in reflecting upon the practice of the radical interpreter in formulating a truth-theory for the language in which such expressions occur (see Davidson, 1984, esp. essays 3, 9, 10, 11, 14, 15 and 16; Evans and McDowell, 1976, intro; McDowell, 1977, 1978). Thinking in this way enables one to give a perfectly adequate account of the relations between thinkers and things which are required if the former are to refer to the latter in thought and talk, whilst resisting any need reductively to formulate this account in terms of certain specific types of causal relations, or relations

⁵I realize that I have said nothing about how exactly testimony might provide such direct demonstrative knowledge about the objects of discussion in certain circumstances. I believe that it does; and I would expect the correct account of this possibility to emerge from an investigation into the way in which what might be called testimonial demonstratives - as, for example, when S says 'that man was a villain' in the context of Evans' case of the pub conversation about Louis XIII - succeed in referring to particular persisting mind-independent objects. This would be very much in accord with the way in which I argue elsewhere that an account can be derived of the way in which perceptual experiences provide non-inferential reasons for empirical beliefs from reflection upon perceptual demonstrative reference to mind-independent things (1999, esp. ch. 6). See Evans (1982, chs. 5 & 9); Fricker (1987); Coady (1992); and McDowell (1994) for important work on the epistemology of testimony.

of any other kind, conceivable quite independently of their role in the intelligible engagement of a rational agent with the world around her.

The relations, then, between a person, in her use of a referring expression, and the thing to which she refers in using it, in virtue of which the expression does indeed refer to that thing, are precisely those relations which prompt an ideal interpreter, in his attempts to make best overall sense of what she says and does, to regard her as talking about that thing in using the expression in question; and there may be no other way to identify or characterize these relations. My proposal is that this process of making sense of what people are thinking and talking about is constrained precisely by considerations of what they have, or could have, knowledge about, most importantly, what it is about which they are provided with demonstratively expressible knowledge by the way in which their attention is focused upon the world in perception, testimony, memory and so on: given, that is, the world as it presents itself to their point of view. Thus, I claim that the relevant process of interpretation is governed by the question which things in the world around them the subjects to be interpreted could have demonstratively-based knowledge about, given the way they are built, the relations in which they stand to such things in using the linguistic expressions which they use in the ways in which they do, and the other ways in which they operate in the world around them.

In other words, there are, amongst the factors determining the interpreter's assignment of a particular object, or natural kind, as the reference of a given referring expression in use in a certain linguistic community, significant epistemic constraints of precisely the kind which my reading of Russell's Principle of Acquaintance, (A) above, demands. The sort of engagement between the language-users and the particular object or kind in question which is required if this assignment is really to make best sense of what they say and do is precisely that involved in their acquiring knowledge about just that

object, or that kind, and retaining this to some extent in memory, in the circumstances of their use of the expression. This is my principle (A): reference to certain objects and kinds requires demonstratively-based knowledge about them. Put another way, the claim is that a person succeeds in singling out a determinate particular or kind in thought, in each of the wide variety of modes of reference by which this is possible, only in virtue of his standing in certain relations with that object or kind which provide him with demonstratively-based knowledge about it. These epistemic constraints upon reference are what, as McDowell puts it, “anchor” the correct semantic theory for a given language to the actual facts of its use (1977, pp. 183-184).

What remains to be shown is exactly how this basis for content externalism, in the Russellian thesis (A), undermines its purported incompatibility with any adequate account of self-knowledge.

Suppose that a person’s belief that p comprises an externalist concept C. The challenging inference would then be this.

- (e1) I believe that p.**
- (e2) If x believes that p, then x’s environment contains (or did contain) C.**
- ∴ (e3) My environment contains (or did contain) C.**

On my view, as I say, the externalist requirement upon possession of the concept C derives from the fact that its semantic value is necessarily a natural kind, say, about which any person who has the concept has demonstratively-based knowledge, where what this amounts to is either that he has knowledge expressible using demonstrative reference to that kind - e.g. ‘that is water in the glass over there’, or ‘that water looks refreshing’ - or that he had such knowledge, some of which is retained, perhaps

linguistically categorized and no longer demonstratively expressible, in memory - e.g. cool water is refreshing to drink', or 'the water in Hinksey pool was very cold'. Now, if the inference set out above is to be an unwarrantedly non-empirical source of knowledge, then its premises must at least be true. The truth of (e1) depends upon the subject's grasp of the content 'p', though, which in turn depends upon his possession of the concept C. From (A), he therefore has demonstratively-based knowledge about C, along the lines suggested above. Hence he is already in a position to arrive at the knowledge that there is (or was) C in his environment if only he turns his mind to the matter. For example, if he already knows that that is water in the glass over there, or that the water in Hinksey pool was very cold, his knowledge that there is water in his environment cannot possibly be any mystery. Therefore this argument cannot possibly constitute a problematic non-empirical source of new empirical knowledge: if its premises are simply true, then the subject already has the wherewithal to arrive at knowledge of its conclusion.

Put slightly differently, the key claim is that the opponent of externalism wrongly neglects the empirical-epistemic constraints upon concept possession which essentially enter into a person's knowledge of (e1) through their application simply to its truth. This already presupposes his standing in an epistemic relation - Russell's acquaintance, as it were - with samples of C. It is the first move in the objector's reasoning at the outset, then, that any adequate account of a person's knowledge of the contents of his own beliefs entails that this is wholly non-empirical, which is to be rejected. For this self-knowledge requires his grasp of the contents of the beliefs in question, his possession of whose component concepts in turn depends upon his empirical-epistemic relations with their semantic values. **Thus, content externalism of this kind is perfectly compatible with an adequate account of a person's knowledge of the contents of his own beliefs.**

To repeat, determinate concept possession is an epistemic skill. It is a matter of a person's being in relations with the relevant worldly semantic values which provide him with demonstratively-based knowledge about such things. This is the source of the externalist requirements upon concept possession. That is to say, the world-involving causal-perceptual relations between a person and certain things in the world around him which are essential to his possession of concepts with those things as their determinate semantic values are precisely the reason-giving relations which undermine the purported difficulty with which I began for combining content externalism with a plausible account of self-knowledge.

Finally, what are the consequences of all of this for **Burge's**, and others', **anti-sceptical strategy**? It faces the following dilemma. Either the causal-perceptual relations governing the content externalism of (B2) are reason-given relations or they are not. If they are, then the transcendental argument is unnecessary. For, as above, the truth of the first premise depends upon the subject's actually standing in such relations with particular physical entities in the world around her, which already provide her with reasons for certain beliefs about them. She has no need for the detour via Burge's argument. This is precisely why the arguments I have been considering pose no threat to externalism. For they do not constitute a non-empirical source of new empirical knowledge: subjects who can run them already effectively have it. If, on the other hand, the relevant causal-perceptual relations are not reason-giving, then the transcendental argument is completely ineffective. For, as I remarked earlier, the conceptual redeployment routine to which Burge appeals in explaining the subject's knowledge of (B1) delivers only her knowledge that the sentence "I am thinking that there are physical entities" is true, in a sense which is neutral between the cases, (a), in which 'physical entities', on her lips, refers to genuine physical entities, because they are what causally explain her use of the relevant expressions, and (b), in which 'physical entities', on her

lips, refers instead to the deceptive intentions of the sceptic's malicious demon, say, or the brain stimulations of the mad scientist. Although she may therefore derive, by Burge's argument, the conclusion that she stands in causal-perceptual relations with those things, whichever they are, to which her use of expressions for physical entities is her normal response, she hasn't the slightest idea what kinds of things these are: in particular, whether they are genuine physical entities or rather some kind of sceptical surrogates for these. Either way, then, the anti-sceptical strategy fails.⁶

References

- Boghossian, P. 1989. 'Content and Self-Knowledge'. Philosophical Topics, **17**, 5-26.
- 1997. 'What the Externalist Can Know A Priori'. Proceedings of the Aristotelian Society, **97**, 161-175.
- Brewer, B. 1999. Perception and Reason. Oxford: Oxford University Press.
- Burge, T. 1979. 'Individualism and the Mental'. In French et al (eds.), Midwest Studies in Philosophy, **4**, 73-121.
- 1986. 'Individualism and Psychology'. Philosophical Review, **95**, 3-45.
- 1996. 'Our Entitlement to Self-Knowledge'. Proceedings of the Aristotelian Society, **96**, 91-116.
- Coady, C. A. J. 1992. Testimony. Oxford: Oxford University Press.
- Davidson, D. 1984. Inquiries into Truth and Interpretation. Oxford: Oxford University Press.

⁶This paper inherits many acknowledgments from my forthcoming book and has acquired its own too. I should like to thank the following for their very helpful comments on earlier versions of this material. John Campbell, David Charles, Bill Child, Martin Davies, Naomi Eilan, Jennifer Hornsby, Brian McLaughlin, Mike Martin, Christopher Peacocke, Johannes Roessler, Helen Steward, Tom Stoneham, Rowland Stout, Ralph Walker and Tim Williamson.

- Davies, M. 1991. 'Individualism and Perceptual Content'. Mind, **100**, 461-484.
- 1992. 'Perceptual Content and Local Supervenience'. Proceedings of the Aristotelian Society, **92**, 21-45.
- 1997. 'Externalism, Architecturalism and Epistemic Warrant'. In C. MacDonald, B. Smith and C. Wright (eds.), Knowing Our Own Minds. Oxford: Oxford University Press.
- Descartes, R. 1986. Meditations on First Philosophy, tr. J. Cottingham. Cambridge: Cambridge University Press.
- Evans, G. 1982. The Varieties of Reference. Oxford: Oxford University Press.
- 1985a. 'The Causal Theory of Names'. In his Collected Papers. Oxford: Oxford University Press.
- 1985b. 'Understanding Demonstratives'. In his Collected Papers. Oxford: Oxford University Press.
- Evans, G. and McDowell, J. (eds.). Truth and Meaning. Oxford: Oxford University Press.
- Frege, G. 1993. 'On Sense and Reference'. In A. Moore (ed.), Meaning and Reference. Oxford: Oxford University Press.
- Fricker, E. 1987. 'The Epistemology of Testimony'. Proceedings of the Aristotelian Society, Supplementary Volume, **61**, 57-83.
- Kant, I. 1929. Critique of Pure Reason, tr. N. Kemp Smith, London: Macmillan.
- Kripke, S. 1980. Naming and Necessity. Oxford: Blackwell.
- McDowell, J. 1977. 'On the Sense and Reference of a Proper Name'. Mind, **86**, 159-185.
- 1978. 'Physicalism and Primitive Denotation: Field on Tarski'. Erkenntnis, **13**, 131-152.
- 1994. 'Knowledge by Hearsay'. In B. K. Matilal and A Chakrabarti (eds.), Knowing from Words. Amsterdam: Kluwer.
- McGinn, C. 1989. Mental Content. Oxford: Blackwell.

- Malcolm, N. 1991. "I Believe that p". In E. LePore and R. Van Gulick (eds.), John Searle and his Critics. Oxford: Blackwell.
- Peacocke, C. 1988. 'The Limits of Intelligibility: A Post-Verificationist Proposal'. Philosophical Review, **97**, 463-496.
- 1993. 'Externalist Explanation'. Proceedings of the Aristotelian Society, **93**, 203-230.
- 1996. 'Entitlement, Self-Knowledge and Conceptual Redeployment'. Proceedings of the Aristotelian Society, **96**, 117-158.
- Pettit, P. and McDowell, J. (eds.). 1986. Subject, Thought and Context. Oxford: Oxford University Press.
- Putnam, H. 1975. 'The Meaning of "Meaning"'. In his Mind, Language and Reality. Cambridge: Cambridge University Press.
- Russell, B. 1917. 'Knowledge by Acquaintance and Knowledge by Description'. In his Mysticism and Logic. London: Allen and Unwin.
- Strawson, P. F. 1959. Individuals. London: Methuen.
- 1966. The Bounds of Sense. London: Methuen.
- 1974. 'Self, Mind and Body'. In his Freedom and Resentment and Other Essays. London: Methuen.
- Wittgenstein. 1958. Philosophical Investigations, tr. G. E. M. Anscombe. Oxford: Blackwell.
- 1975. On Certainty, ed. G. E. M. Anscombe and G. H. von Wright, tr. D. Paul and G. E. M. Anscombe. Oxford: Blackwell.
- 1980. Remarks on the Philosophy of Psychology, vol. I, ed. G. E. M. Anscombe and G. H. von Wright, tr. G. E. M. Anscombe.

SELF-KNOWLEDGE AND EXTERNALISM

BILL BREWER

- (E1) I believe that p.
(E2) If \underline{x} believes that p, then \underline{x} 's environment is thus-and-so.
∴ (E3) My environment is thus-and-so.
- (w1) I believe that water is wet.
(w2) If \underline{x} believes that water is wet, then \underline{x} 's environment contains (or did contain) water.
∴ (w3) My environment contains (or did contain) water.
1. Adequate acct. of self-knowledge → (E1) NEK (non-empirically knowable).
 2. Content externalism → (E2) NEK.
 3. So content externalism plus an adequate acct. of self-knowledge → (E3) NEK.
 4. Non-empirical knowledge of empirical facts like this is intuitively intolerable.
 5. Thus, content externalism is incompatible with an adequate acct. of s-k.
- (B1) I think that there are physical entities.
(B2) If \underline{x} thinks that there are physical entities, then \underline{x} bears causal-perceptual relations to physical entities.
∴ (B3) I bear causal-perceptual relations to physical entities.

Solution: There is an essential empirical component in self-knowledge.
For determinate concept possession is an epistemic skill.

Alternatives

1. Content externalism is not committed to the a priori knowability of specific instances of (E2).
2. Knowledge of specific instances of (E3) cannot be acquired on the basis of any argument of this form.
3. There is no difficulty in the idea of wholly non-empirical knowledge of empirical matters of fact.
4. Self-ascriptions of beliefs are not, contrary to appearances, genuinely truth-evaluable statements.

Set 3 and 4 aside; consider and reject 1 and 2.

(d1) I believe that that ball will go into the pocket.

(d2) If \underline{x} believes that that ball will go into the pocket, then \underline{x} 's environment contains that ball.

\therefore (d3) My environment contains that ball.

(LP) Epistemic warrant cannot be transferred from A to B, even given an a priori known entailment from A to B, if the truth of proposition B is a precondition of the knower being able to believe the proposition A.

(M1) r_1 and r_2 are real numbers

(M2) The product of any two real numbers is a real number

\therefore (M3) $r_1 r_2$ is a real number

(A) A person's capacity to make determinate reference to certain objects and kinds in belief depends upon his having demonstratively-based knowledge about them.

(e1) I believe that \underline{p} .

(e2) If \underline{x} believes that \underline{p} , then \underline{x} 's environment contains (or did contain) \underline{C} .

\therefore (e3) My environment contains (or did contain) \underline{C} .

Thus, (A) resolves the prima facie tension between content externalism and self-knowledge.

And Burge's anti-sceptical strategy fails.

SKE.B

1. Thesis: externalism goes only as far as it applies to:
the subject's experiential point of view.
2. This surely applies to egocentric spatial concepts.

Left and right are that direction, and that direction. One cannot characterize them other than by physical demonstratives in this way. Yet they are essential constituents of our experiential point of view.

3. Thus softened up, I'm inclined to extend the view to particulars: this lectern, and that chair, say.

On this account, the metaphor of a subject's experiential point of view is to be taken entirely literally. My experiential condition just is my having this view of the world from here now in these conditions.

This is a view you can share - as the guide-books always insist when they say: "the view from the Marin headlands is spectacular; you must not miss it".

And such a view, my conscious experience, has as constituents the particular objects upon which it is a view: this lectern, and that chair, say, just as the Golden Gate is a key feature of the view from the Marin headlands.

This is so even though I cannot infallibly distinguish my current experiential condition from the view I might have had of a different lectern, different chairs etc.

The view I shared with my friend a couple of weeks ago from the Marin headlands is a distinct view from any infallibly indistinguishable duplicate some millionaire might construct in Texas,

precisely because it is constituted by distinct objects: the actual Golden Gate, rather than any qualitatively identical but numerically distinct duplicate.

This is how I think we should think about the subjective experiential points of view of people upon the world around them.

4. Having got that far:

insofar as externalism can be pushed further, from particular objects to natural kinds, say, then I claim that this externalization of the subjective point of view must be pushed equally far.

Both stop together.

So the arguments I've been considering neither pose a threat to externalism, nor a transcendental defeat of scepticism.

5. I take my disagreement with Searle to be over where the externalism stops.

He thinks it stops so soon it never gets started - although I don't know what he thinks about space.

Against his descriptive account of perceptual demonstratives I say.

(i) Odd acct. for someone who's motivation for stopping the externalism so soon is presumably that we can't be mistaken about how things are for us subjectively. For I certainly don't think I'm thinking "the lectern causally responsible for this experience" when I think about this lectern.

(ii) I have worries about the coherence of the internal demonstratives: how do they work?

(iii) Does this actually add up to thought about mind-independent things at all?

What conception can the subject possibly have of the thing causally responsible for this experience?

Either this is a construct out of the experience itself, and others; in which case he certainly knows what kind of thing it is, but it is hardly objective.

Or it is not, in which case it might, for all he knows, be some kind of Kantian thing in itself, or Berkeleyian idea in the mind of God, or what-have-you.

Either way, we do not seem to have an account of thought, with understanding, about the mind-independent world of particular macroscopic objects in space and time.

(iv) For what it's worth, this last point seems to me very closely related to the stuff Cora Diamond was saying about a kind of PLA in the Tractatus objection to Russell's theory of other minds as known only by description.