

Analytic Transcendental Arguments

Jonathan Bennett

From: P. Bieri et al. (eds), *Transcendental Arguments and Science* (Reidel: Dordrecht, 1979), pp. 45–64.

1. Locke on the objective world

Someone who thinks that his own inner states are the basis for all his other knowledge and beliefs may wonder how anything can be securely built on this foundation. He need not actually doubt that his own edifice is securely founded, though he may pretend to have doubts about this in order to consider how they could be resolved if they did occur.

This person is a ‘Cartesian sceptic’. which implies that he is not sceptical at all. He is untouched by such crude English moves as Locke’s protest that ‘nobody can in earnest be so sceptical’, or Moore’s holding up his hand as proof that there is a physical object. Such intellectual bullying is irrelevant to a serious inquiry into how epistemic foundations relate to epistemic superstructure.

Locke’s answer was that ‘There is an outer world’ is a good explanatory hypothesis: various facts about my inner states, he thought, are best explained by the theory that there is an outer world. For example, from the fact that some of my ‘ideas’ occur without my willing them to do so, Locke infers that ‘it must needs be some exterior cause. . . that produces those ideas in my mind’ (IV.xi.5). Let us concede for purposes of argument that if some of my ‘ideas’ are involuntary then something other than myself exists; still,

nothing could follow about *what* exists other than myself. Can we strengthen the conclusion by strengthening the premises? Could a more contentful belief about an outer world be defended as explaining certain further facts about my inner states, e.g. about the order or regularity which they exhibit? Locke does argue like that, but unfortunately he pollutes all his premises—which should be purely about inner states—with an admixture of statements about the outer world; for instance, he uses the premise that men with no eyes have no visual states. But that seems to be an accidental defect in Locke’s treatment. He could have cleansed his premises, as Hume nearly did, so that they spoke only of the order, coherence etc. of one’s inner states; and I am sure that if he had done so he would still have argued that those facts are best explained by the hypothesis that there is a world of physical objects. Could such an argument be valid?

That depends upon how it handled the following problem. If one is to argue in the form: ‘On the strength of data D, we are entitled to accept theory T, because T is the best explanation for D’, one must have an account of what makes this explanation better than some other. Why, for instance, is the physical-world explanation for the orderliness of my inner states better than a Cartesian-demon one?

Sometimes one explanation is superior to another because there are facts which show it to be more plausible, more likely, more in tune with reality, than the other. There is a sudden loud noise, and we wonder ‘Was that thunder or an explosion?’; and someone who knows the region well can tell us which explanation is more likely. But I could not have that sort of reason for preferring the outer-world theory to a rival explanation of my inner states, because there is nowhere for the required supporting facts to be found. They cannot, without begging the question, be sought in the outer world; and since they must show how my inner states relate to something else they cannot be found purely within my inner states either. And the outer and inner exhaust the territory. This is quite different from the problem of explaining one particular noise, or even all noises of a given kind, or for that matter all noises whatsoever: each of those problems is limited, leaving plenty of territory not covered by the problem and thus available as a source of helpful information—we have access to the causes of noises otherwise than through hearing them. Whereas the problem of explaining the orderliness of my inner states is so broad that it sprawls over all the possibly relevant territory.

Locke seems not to have noticed this difficulty about what justifies one in preferring one explanation to another. (Sometimes he writes as though the outer-world theory were the only possible explanation for the facts about inner states: but that is a rhetorical extravagance.) Nothing in his discussion rules out the idea that the outer-world theory is the best because it is the most plausible, the one with the most independent factual support; and so even if he had cleansed his premises of their question-begging elements Locke would not, I think, have been on the track of a viable solution to the outer-world problem. I now leave Locke, and stay with his problem.

2. Another merit in explanatory theories

One theory can be superior to another in being simpler or more powerful or both, and that basis for comparing two theories does not appeal to any further range of relevant fact. If we don’t have access to any such further range, it seems that we must base our theory-preferences upon considerations of power and simplicity. It also seems that in such a case the only ground we can have for accepting a given theory is just that it is superior in power and economy to any known rival explanation of the same data.

That, I take it, is an ingredient in a widely accepted account of what entitles one to accept high-level theories in physics. A restricted item of high-level theory may be defended on the ground that certain facts make it likely or plausible, but those ‘facts’ must involve other bits of high-level theory which are being taken for granted. If what is in question is the whole level of theory which pertains to sub-atomic items, say, the only way we can defend it is by showing how it helps us with lower-level data, and especially how it facilitates lower-level predictions. Along with the high-level theory we accept many conditionals running upwards and downwards between the two levels; this creates high-level inferential routes from some low-level statements to others, and these routes can be useful, to put it mildly. For instance, someone who judges that the H-bomb he has made is in good working-order is moving from some observed facts to a prediction about further observed facts, and he couldn’t easily do this except by going through abstract physical theory. And that, it seems, is our whole justification for accepting that level of theory—namely that it organizes the lower levels for us better than does any known rival.

Does that make the high-level theory genuinely explanatory? Well, it is explanatory in the sense that it confers conceptual unity on the lower level, and thus facilitates predictions, eases memory, and produces intellectual satisfaction. Apparently that is as much explanatoriness as it could possibly have. But it is natural to want something more: a genuinely explanatory theory, one is inclined to think, must be not merely useful but true. The contrast between utility and truth, echoing the quarrel between ‘instrumentalism’ and ‘realism’, should be discussed a little.

Let T be some theory which does a good job in conceptually organizing certain data, and indeed a better job than any known rival theory. The thought that T may nevertheless not be *true* might be cashed out in any of three ways. **(1)** Perhaps logical space contains a theory which organizes past and present data even better than T does, but which we haven’t thought of. **(2)** Perhaps there are available data—data which we would have if we acted in certain ways—which, when added to all the data we do have, would show T to be inferior to some rival in power or simplicity or both. **(3)** Perhaps in future we shall have data which, when added to the data we now have, will show T to be inferior to some rival. The idea that a merely useful theory may not be true could be fed by any of those sources, i.e. by an awareness that a superior rival might be found through intellectual effort, perhaps in combination with experimental effort or sheer passage of time. There can be no quarrel with a ‘realism’ which reminds us of these possibilities, and subjects every currently favoured theory to the ‘sceptical’ thought that it might be dislodged by a superior rival.

But there remains the stronger kind of ‘realism’ which says that even if theory T is in fact superior to any rival which could be devised, in relation to any data which will or could be encountered, there is still a question as to whether

the theory is true. This involves the thought: ‘Perhaps the real principles which govern reality are somehow less simple and sweeping than are the regularities which appear to us or which we could encounter by letting time pass and by experimental effort.’ Someone who asks us to take that thought seriously should explain what he means by a theory’s being ‘true’, or how he takes ‘the principles which really govern the universe’ to be conceptually different from ‘the regularities which are or could be encountered in our experience of the universe’. Since I don’t think that he can coherently explain these things, I think that this strong ‘realism’, and its associated ‘scepticism’, are incoherent. But I shan’t insist upon that.

3. Verificationism

I use ‘verificationism’ as a label for a certain view about my entitlement to believe that there is an objective realm. One of its components says that I am entitled to accept that theory because the latter is superior to any known rival in its power to confer conceptual unity on my inner states. For example, the theory which says simply that my inner states are caused by a demon whose pleasure it is to give me false beliefs, though very simple, has almost no power to unify, organize, generate predictions and so on. There could be a much more powerful demon theory, because at any time when T is my actual theory of the outer world there is a rival theory which says ‘My inner states are and always will be caused by an undeterrable demon whose pleasure it is to make me believe T’. That will explain everything that T explains; but it takes slightly longer to do so, and brings no compensating advantages. Of course I might come to be in states which were less satisfactorily handled by T than by its demonic cousin; and the verificationism I am presenting does not rule out that possibility, any more than it does the

possibility that even now there may be some superior rival to T which I haven't thought of.

The other component in verificationism is its ruling out of the sort of scepticism which is generated by what I shall call 'transcendental realism'. The latter is the view that the truth or falsity of T transcends all the facts about inner states which I do, will or could have; so that even if we assume that nothing I can do would produce states of myself which will be better handled by some other theory than by T, the question 'But is T actually true?' remains unanswered. According to verificationism, this sort of transcendental realism involves a conceptual mistake about what it is for a theory to be true, and the associated scepticism should be condemned as the asking of a non-question.

What verificationism says about the basic merits of the outer-world theory is something which Hume came close to before being swept away from it by cross-currents in his thought (pp. 195–7); and it seems to me obviously correct. But it is less obvious that verificationism is right in claiming that it has described all the merits that a theory could possibly have, i.e. its rejection of the transcendental realist thought that the truth of the outer-world theory is not secured by the theory's having no possible rivals which surpass it in simplicity or power.

This extra element in verificationism is something it shares with phenomenalism; but I am envisaging verificationism as being cautious enough to avoid the fatal defects in phenomenalism. In particular, I take it to be saying that a proper understanding of the meanings of objectivity-statements involves grasping the kinds of ways in which they can be related to statements about inner states, through conditionals running one way and conditionals running the other; but not as associating the meaning of an objectivity-statement with any definite set of conditionals,

thus freezing its meaning in an objectionable manner; and still less as trying to relate the outer and the inner through biconditionals, as phenomenalism does.

Those cautiously negative remarks about verificationism imply that it is, in my hands at least, a rather sketchy doctrine. Still, it has enough content to be non-trivial, while also being cautious enough to be possibly true, I believe that it is true, as far as it goes, which puts me on the same side as Kant: see, for instance, his remark about 'that unity' in our items of knowledge 'which constitutes the concept of an object', and what he says about 'the concept of body' as 'a rule for our intuitions' (A 104–6). Verificationism is not, of course, the whole content of Kant's 'transcendental idealism', but it is an important part of it. And I take it that when Kant speaks of regarding the world as 'a thing in itself' at least part of what he means is: regarding the world as something whose real nature is not exhausted by all the facts about actual and achievable experience.

But I shall not argue from verificationism taken as a premise, nor shall I attach much weight to my view that it was accepted by Kant. All I need is to mention 'verificationism' and have you know what I mean; and even that will not come until Section 8 below.

4. Transcendental arguments

The inquirer or 'sceptic' with whom I began has no problems over his knowledge of his inner states, but he does have one over his beliefs about the outer world. That was indeed how Descartes thought he was situated: he apparently couldn't grasp the idea that self-knowledge might also need explaining; When Gassendi implicitly challenged him to explain how "you manifest yourself to yourself merely by the operation called thought", Descartes lost his temper (pp. 716, 799). Nor was the question taken any more seriously by the

British empiricists; and although Spinoza and Leibniz had views about it, they are not relevant to my main topic. The first relevant contribution came from Kant, who argued for, and from, statements about what is required for awareness of one's inner states; and, most notably, presented reasons for thinking that if the sceptic knows what he thinks he does then he must also know some of the things which he thinks he doesn't. And in our own time Wittgenstein, Strawson, Shoemaker, Rorty and others have reached fairly Kantian conclusions by means of fairly Kantian arguments.

The last phrase needs comment. The philosophers in question have offered what could be called analytic arguments, by which I mean ones built entirely out of the materials of conceptual analysis and conceptual inter-linkage, without taking anything from what Strawson has called 'the dark side' of Kant's thought—the side that entertains the notion of things as they are in themselves, implies that the mind atemporally creates its objects, and so on. I believe that one can indeed build a significantly Kantian argument without drawing on 'the dark side', but I shan't defend this view (e.g. against Hintikka and Rosenberg). The organizers of this conference asked me to discuss a certain kind of conceptual argument, and my concern is with that kind of argument—its powers and its limitations, but not its historical antecedents. When I call such arguments 'transcendental' that could be taken as a mere label with no Kantian implications at all.

I take a 'transcendental' argument to be one which aims to rebut some form of scepticism by proving something about the necessary conditions for self-knowledge, self-consciousness, or the like. That is a narrower usage than some writers have adopted. Any argument which I count as 'transcendental' will imply that the scepticism which it opposes is in some way self-defeating: but I don't label as

'transcendental' every argument which seeks to show that some kind of scepticism is self-defeating—e.g. Malcolm's argument that scepticism about freedom is self-defeating because if there is no freedom there are no intentions, and thus no assertions, so that no-one can truthfully assert 'There is no freedom'. Nor do I count as 'transcendental' every argument 'to the conclusion that the truth of some principle is necessary to the possibility of the successful employment of a specified sphere of discourse' (Griffiths, p. 167)—a usage which makes the term 'transcendental argument' very broad indeed (as is pointed out by MacIntosh, pp. 185–6). An argument about what is required for the possibility of language as such is a plausible candidate for the 'transcendental' label (Kekes), and I shall sometime mention such arguments in passing. But I wouldn't unqualifiedly call them 'transcendental arguments' unless they included the claim (which I doubt) that self-knowledge conceptually requires linguistic capacity.

On the other hand, I use 'transcendental argument' more broadly than Gram does: since 'Kant has the only clear historical title to [the] notion' of a transcendental argument, Gram won't count as 'transcendental' any argument which lacks the highly 'peculiar character' demanded by some remarks of Kant's; but this leads him to conclude that there cannot be any valid transcendental arguments (p. 15 and *passim*). Körner also reaches this negative conclusion on the basis of a Kantian understanding of what a 'transcendental argument' is.

I have no quarrel with any of these writers. I merely use 'transcendental argument' in my own way.

5. Intuitions, concepts, judgments

Now, the Cartesian inquirer thinks that he knows what his inner states are while having a problem about the outer

world; to which a transcendental argument replies that if he has self-knowledge then this entails some partial solution to the outer-world problem. Entails what? Well, one might try to argue that if someone has self-knowledge then:

- (1) his inner states must be thus and so; or
- (2) he must have such and such concepts; or
- (3) he must employ such and such concepts.

I take it that (3) involves the other two: to employ a concept you must possess it and must have something to apply it to. Can we stop short of (3)? Could it be shown, for example, that self-knowledge requires inner states to which objectivity-concepts *could be* applied, without requiring that they actually be applied? I doubt it, and I have never seen a transcendental argument which aimed at (1) but not (3).

But some have tried to prove (2) without (3) and perhaps also without (1). That is, they have tried to prove only that self-knowledge requires a disposition to apply certain concepts if one should have the appropriate data. Strawson has an argument which is like that. (It is a 'transcendental argument' in my sense, though it concerns not the whole outer world but only other minds.) He argues that self-knowledge requires one to have a concept of 'person' which can be applied on the strength of behavioural indications, but he does not argue that the self-knower must have any actual opportunities to apply this concept to anyone other than himself (Strawson, 1959, ch. 3). This is not the place to criticize Strawson's argument in detail; but I have a general reason for preferring to argue for (3) rather than for (1) or (2) alone.

The reason stems from Kant's insight that self-knowledge involves intellectual activity: to know what one's inner states are like is to make judgments of certain kinds. Now, just as one might show that to climb a mountain one must bend one's knees, or that to balance the books one must add

figures, so we may be able to show that to do whatever is involved in self-knowledge one must do certain ancillary things; and since in this context the relevant 'doing' is judging, such an argument would conclude that self-knowledge requires one to make certain sorts of judgments, i.e. employ certain concepts. That would lead to a type-(3) conclusion. To argue only for (2), one would have to show that in order to do what is involved in self-knowledge one must be intellectually equipped to do certain other things without necessarily doing them; and I don't see how such an argument could succeed.

6. A simple transcendental argument

As a basis for further discussion, I shall now present an actual transcendental argument whose conclusion is of type (3). The argument is my own, though I found it while thinking about Kant's Refutation of Idealism (Bennett §51).

It concerns beliefs about one's past inner states. Starting with someone who has beliefs about his present states, I argue that he cannot add beliefs about his past unless he brings some of his inner states under objectivity-concepts, i.e. unless he reasonably regards himself as inhabiting an objective realm.

Prima facie, it seems that the protagonist's states might include recollections of his past states, thus giving him access to his past even if all of his inner states are just a chaotic jumble which cannot be treated as perceptual contact with an outer world. The first step in my argument is to show why that is not really possible.

Adopt the hypothesis that the protagonist does have recollections of his past inner states. That implies that a sub-set of his inner states contain representations of other inner states: as well as pains and buzzings and glows he also has what I shall call K-type representations of pains and buzzings and glows; and we are supposing these to be his recollections

of earlier pains etc. But this account contains no basis for connecting a K-type representation with the protagonist's past. We are pretending that when he judges 'I have a K-type representation of state S' he can infer the judgment 'I was earlier in state S'; but this supposed inference from the K-type representation to a judgment about the past is simply idle for our protagonist—it is, in Wittgenstein's metaphor, a wheel which turns though nothing turns with it. I say this for two reasons.

Firstly, our protagonist has no way of using any judgment about the past: since his inner states are treated by him as a chaos, in that he doesn't bring them under any general principles of order, he cannot use the judgment 'I was earlier in state S' as evidence for any other judgment. This might be turned into an attack on the assumption that he even makes judgments about his present states; but I shan't press that point, because I want to grant him his present-tense judgments so as to show that he can't have past-tense ones.

Secondly, and more important in my argument: there is a one-to-one correlation, for the protagonist, between judgments about the past and the present-tense judgments on which they are based: he is in a position to judge 'I was in state S' when and only when he is in a position to judge 'I have a K-type representation of state S'. This means that the supposed addition to his *Weltanschauung* of the supposed judgments about the past is a routine, simple, mechanical business; something which does not introduce any complexity, or any further element of structure, into the protagonist's intellectual situation.

Compare our protagonist with someone who, given the very same inner states, takes his K-type states to be representations of future states. There is no real content to this supposed difference between the two: the judgments of one can be systematically paired off against those of the other;

and the illusion that there is a real difference arises purely from my having chosen—with no basis in the given facts—to word the judgments differently.

Then compare those two with a third person who, given the same inner states, does not regard his K-type states as representations at all. He notes that he is sometimes in pain and sometimes in a state which is R-related to pain, sometimes experiences buzzing and sometimes experiences something which is R-related to buzzing, and so on; but he does not take R to be a representing relation, let alone take it to be specifically the relation of recollecting or foreseeing. This too sounds like a totally different story, but only because I have chosen to word it differently: nothing in the real content of the story justifies one wording rather than another.

I conclude that if the protagonist's inner states are treated by him as a chaos, he can have no working concept of the past, and thus no concept of the past.

Now suppose that he reasonably believes that he has experience of an outer world: his inner states are suitable, and he has the appropriate concepts and applies them in objectivity-judgments. Obviously, he can now draw conclusions from some of his judgments about the past; but my main concern is not with what flows *from* those judgments but with what flows *into* them, i.e. with the protagonist's grounds for making judgments about the past.

He must still base those judgments upon present data—his present inner states—but he is no longer restricted to one datum per judgment, for he can now bring several of his present inner states to bear on a single judgment about his past. This can happen in three ways. His judgment that he was in state S at past time t_1 maybe confirmed or disconfirmed (a) by his recollections of his states at other times, as when his recollection of falling at time t_0 to confirms his judgment that he was in pain at t_1 ; or (b)

by his recollections of his other states at t_1 , as when his recollection of seeing the sun at t_1 confirms his judgment that he felt warm at t_1 ; or (c) by data other than recollections, as when his seeing of ashes at t_2 confirms his judgment that he saw a fire at t_1 . Each of these involves general judgments as well—falls are followed by pain, sunshine is accompanied by warmth, ashes are preceded by fire—and the acceptance of such ‘laws’ is of the essence of a belief in an outer world.

My point is not that a judgment is more secure if several data confirm it. The argument does not concern the *trustworthiness* of judgments about the past but rather their *possibility*. Now that several of the protagonist’s states may bear upon a single one of his judgments about the past—confirming or disconfirming it—he has a complex relationship between ‘I was. . .’ and ‘I have a K-type representation of being. . .’, rather than a simple one-to-one mapping, so that his concept of how he was in the past is no longer idle. Each of his judgments about how he was at t is a summing-up of where the balance of the evidence lies—an over-all judgment based on what *seems in most ways* to have been the case at t . When he had no use for objectivity-concepts, on the other hand, there was at most one way in which something could seem to have been the case at t ; and so the notion of ‘what was the case’ collapsed into that of ‘what seems to have been the case’, which in turn degenerated into something which had nothing at all to do with the past.

That completes my ‘transcendental argument’ for the view that anyone who has beliefs about his own past inner states must also apply objectivity-concepts.

7. The pursuit of strength

When arguing for a conditional, one wants its antecedent to be as weak, and its consequent as strong, as possible.

My antecedent is ‘If someone has beliefs about his own past inner states. . .’: I have not refuted the view that someone could at each moment know what his present inner states are, while having no beliefs about his past states or about an outer world. Still, that may be refutable by further argument (and if it isn’t then my argument in Section 6 is not after all a ‘transcendental’ one in my sense). It is plausible to suppose **(a)** that beliefs are impossible unless one has standards for deciding whether they are true or false (Wittgenstein, §258; Rorty (1970), p. 222; Harrison, pp. 56–57), and **(b)** that I can have standards for my present application of a concept only if I have beliefs about my past applications of it (Kant, A 101; Wittgenstein, §260). Those two theses would jointly strengthen the conclusion of my argument by weakening its antecedent to ‘If someone has beliefs about his inner states. . .’. I’m not sure that this will go through, because I am not certain of **(b)**; but all I want here is to indicate some possible argumentative strategies.

I have to admit that in the conclusion of my argument, the consequent needs to be weakened. I have said that if someone makes judgments about the past he must make objectivity-judgments, but really all my argument shows is that he must bring his inner states under a system of lawlike generalizations which lets him connect several judgments about the present with a single one about the past. A system of laws which implied the existence of an outer world would secure this result, but perhaps some weaker system might also suffice. On this question, I am defeated. I have tried to invent laws strong enough to put the concept of the past to work yet too weak to introduce objectivity-concepts; and I have tried to prove that there cannot be such laws; and I have failed in both attempts. All I can say is that a non-idle concept of the past needs a kind of ordered complexity which I cannot see how to get without

also bringing objectivity-concepts into play.

Wilkerson has suggested that it is typical of transcendental arguments that they don't produce necessary conditions for self-knowledge or whatever, but only sufficient conditions together with the claim that 'we are incapable, given our present conceptual resources, of thinking of any other conditions which would be. . . sufficient' (Wilkerson, p. 211; see also Smith, p. 159). There are indeed interesting arguments of that form, though the best example I know is not a transcendental argument in any accepted sense. I allude to Strawson's linking of objectivity with spatiality, by an argument which shows that objectivity involves certain requirements which are fulfilled by spatiality and not by anything else one can think of (Strawson, 1959, ch. 2; Bennett, p. 43). But I see no reason to think that transcendental arguments must be of that inconclusive sort. I think that my argument in Section 6 shows conclusively that someone who has a working concept of his own past must bring his inner states under general laws; and I see no reason in principle why someone should not strengthen that result—still using, conclusive arguments rather than Wilkersonian ones—by weakening the antecedent or strengthening the consequent or both. With any argument which is not rigorously formalized, there is admittedly a marginal possibility that some gap has been overlooked; but I take Wilkerson to be claiming a special inconclusiveness in transcendental arguments as such, and that is what I am questioning.

A famous transcendental argument of Strawson's has a conclusion which is stronger than mine in both ways (Strawson, 1966, pp. 72-112). In contending that if someone has any awareness of his present states he must employ objectivity-concepts, Strawson offers a conditional with a weaker antecedent and a stronger consequent than mine. But even with the help of Rorty's patient reconstruction of

that argument of Strawson's (Rorty, 1970), I am still not sure that I understand it. In contrast with that, my argument in Section 6 is relatively plain and clear; but it achieves this at the price of having a rather weak conclusion, and it *may* be that to strengthen it must one must forgo my kind of plainness in favor of Strawson's kind of subtlety, depth and provocative elusiveness.

8. A dilemma

Even if it could be proved that self-knowledge requires the use of objectivity-concepts, i.e. requires the reasonable belief that there is an outer world, one might still wonder what help this gives with the problem confronting the Cartesian inquirer, the problem which Locke tried to tackle. The question could be expressed thus: even if a self-knower must *believe* that there is an outer world, the question still remains as to whether there *is* such a world. We may have a transcendental argument which 'shows that the referents of the concepts in question are not merely useful fictions. But, for all such an argument shows, they may still be fictions, even if indispensable ones' (Tlumak, p. 263; see also Ayer, pp. 105-9, and Smith, pp. 165-8). But I find that formulation hard to work with: it puts the sceptical inquirer in the position of having to say 'I believe that P, but is it really true that P?'; and that slightly peculiar question raises issues which I prefer not to go into here.

A better formulation of essentially the same issue is provided by a dilemma. Our Cartesian inquirer is asking for the credentials of the theory that there is an outer world; but how does he construe this? Does he **(a)** construe the theory in a verificationist manner, so that its truth could be sufficiently assured by facts about how well it conceptually organizes his inner states? Or does he rather **(b)** construe it in a transcendental realist manner, so that its being more

economical and powerful than any possible rival theory does not entail that it is true? The dilemma is this: if **(a)** he accepts verificationism, then that gives him a reasoned justification for accepting the outer-world theory, and there is no need for any transcendental argument; but if **(b)** he has a transcendental realist understanding of the outer-world theory then transcendental arguments will be powerless to help him with his inquiry. The case for **(a)** is obvious. The case for **(b)** depends on the fact that transcendental arguments can only prove conclusions to the effect that if someone has self-knowledge then *he* must satisfy certain conditions; so an outer world can enter the picture only by inference from propositions about the self-knower. There seems then to be no hope of showing that a self-knower must inhabit an outer world unless 'There is an outer world' is construed in a verificationist manner. Such a construal was clearly essential to the argument which I presented in Section 6 above.

So there is the problem: it seems that **(a)** if a verificationist position is accepted, transcendental arguments are not needed to help the Cartesian inquirer, and **(b)** if it is not accepted they cannot help him. I once discussed **(a)**, but **(b)** seems to have escaped my notice (Bennett, §52); and **(b)** but not **(a)** was implicitly pointed out by Williams in a remark about 'Kant's insistence that his transcendental arguments gave knowledge of how things must be only because things were not things in themselves' (Williams, p. 218). It was Stroud who combined the two points to generate an argument for the conclusion that transcendental arguments cannot do any valid work in answering sceptical inquiries about the well-foundedness of the belief in an objective realm.

Stroud suggests rather than asserts his conclusions; and, strictly, he is addressing himself not to transcendental

arguments generally but just to certain recent and prominent examples. I believe, though, that he could have cast his net wider by construing 'verificationism' more broadly than he does most of the time. I shall try to explain.

The main emphasis in Stroud's paper is on a somewhat narrow kind of verificationism which involves claims of the form 'x makes sense only if P', based on a verification principle which lays down necessary conditions for intelligibility. Each transcendental argument which Stroud discusses does indeed seem to use such a principle, and in each case it is arguable that the relevant kind of scepticism can be met just by the appropriate verification principle, if the latter is correct. So on this basis Stroud can indeed generate a dilemma with which to confront those transcendental arguments. But it hardly makes trouble for the argument which I presented in Section 6 above. That argument relies on a view about what makes a concept honest or non-idle; but it would, I suggest, be strained and inaccurate to call that a 'verification principle'.

However, there is broader way of construing 'verificationism', namely as the view that there is no coherent question about a theory's truth over and above all the questions about its comparative success in conceptually unifying the data which fall under it. My argument in Section 6 is 'verificationist' in that sense: its conclusion was a proposition about how a self-knower (with a concept of his past) must conceptualize his inner states, and so it was about the outer world only on a 'verificationist' construal of statements about the outer world. Furthermore, I don't see how any transcendental argument could be even *prima facie* relevant to sceptical inquiries unless it was 'verificationist' in this broad sense.

The narrower construal of 'verificationism' predominates in Stroud's paper and in some of the subsequent literature. It is, for instance, the only possible basis for bracketing this

work of Stroud's, as Rorty does (1971, p. 4), with Judith Jarvis Thomson's discussion of a private-language argument (Thomson, p. 29). The broader construal of 'verificationism' is also present in Stroud's paper, as is noted by Goldman (pp. 106–7), but its role is a rather quiet and subdued one; and, in particular, it is not used explicitly to generate the broader dilemma which challenges the relevance to scepticism of every transcendental argument. That broader dilemma is now my sole concern.

9. Two counter-attacks against the dilemma

Stine has defended transcendental arguments against the dilemma, by attacking the horn of it which says that such arguments are powerless unless verificationism is presupposed (Stine, pp. 49–51). Rather than presupposing verificationism, Stine says, a good transcendental argument could constitute a defence of it. That, if it were right, would put transcendental arguments right back into service in answering sceptical inquiries; but I don't think that it is right. Stine's argument seems to run as follows. If a transcendental argument proves that every self-knower must bring his inner states under an outer-world theory, this *justifies* the acceptance of the theory; but that is to justify this conceptualization of one's inner states; and isn't that all that verificationism sets out to do? Unfortunately, it is not. To answer the sceptical inquirer, the verificationist has to maintain not merely that a certain way of using objectivity-concepts is all right, but also that a certain other way of using them is conceptually inadmissible—i.e. that when everything has been said about the comparative utility of the outer-world theory no coherent question remains to be asked. No transcendental argument has the power to show that.

Hacker has counter-attacked against the horn of the dilemma which says that if verificationism is presupposed

there is nothing for transcendental arguments to do (Hacker, p. 84). He envisages a sceptic who is not my mild Cartesian inquirer but rather a radical fellow who makes claims about his inner states while denying that anything else even makes sense. 'The existence of anything other than my own states', he says in effect, 'is not just dubious, not just false, but downright unintelligible.' So there he sits, untouched by verificationism; and then a transcendental argument sets him back on his heels by showing that if he does not accord meaning (and indeed truth) to statements about an outer world he cannot know about his own inner states.

I am not sure what exactly Hacker's sceptic is saying. Unless he is direly confused, he must agree that one can give to outer-world statements the sense conferred on them by verificationism and employed in transcendental arguments. And Hacker's sceptic does agree to this, for he sees himself as 'securely in possession of his array of "Cartesian thoughts", out of which he may, so he conjectures, build up a Humean *fiction* of an objective outer world'; but such a fiction would confer meaning on objectivity-statements; so it cannot be that sort of meaning which Hacker's sceptic is denying to them.

The only alternative construal of him that I can find is this: Hacker's 'sceptic' is merely saying that objectivity-statements are unintelligible if they are construed in anything but a verificationist manner, i.e. that they make no sense if the objective realm is taken to be a 'thing in itself'. That, of course, isn't 'scepticism' in any reasonable sense. Also, it is not in conflict with any valid transcendental argument. For a transcendental argument to have any impact on this position—namely the position of someone who says that objectivity-statements have a verificationist sense and no other—it would have to prove that self-knowledge requires that objectivity-statements be accepted both on verification-

ist and on thing-in-itself construals; and I don't believe for a moment that such a thing can be proved. Strawson in his 'Persons' chapter does seem to attempt something of the kind: he argues that self-knowledge requires a preparedness to apply a concept of 'person' in such a way that 'x is a person' is guaranteed by behavioural (verificationist) premises and in turn guarantees mentalistic (thing-in-itself) conclusions (Strawson, 1959, pp. 106–110). But I agree with the majority opinion that this argument of Strawson's does not succeed.

10. Unradical scepticism

I see no prospect for doing any damage to the horn of the dilemma which says that transcendental arguments are powerless if verificationism is not presupposed. If someone asks 'Is the outer-world theory true?', and construes this in a non-verificationist manner as an inquiry about the world as a thing in itself, then no transcendental argument can either answer it or show that it ought not to be asked. Where the truth of verificationism is denied or questioned, transcendental arguments must stand mute in face of all sceptical inquiries.

However, I do think there is something wrong with the horn of the dilemma which says that if verificationism is assumed then *it* will answer the sceptical inquiries, so far as they can be answered, leaving transcendental arguments idle. To that extent I agree with Hacker, though not with his account of what it is that transcendental arguments can achieve which mere verificationism cannot. (In all of this, I am considering only their power to answer sceptical inquiries. The exploration of the conceptual requirements for self-knowledge, considered just as a part of the philosophy of mind, lies beyond my present scope.)

Suppose we are confronted by a 'sceptical' inquirer who concedes what verificationism demands: that is, his question

about the outer-world theory is only the question as to how well that theory functions in comparison with possible rivals. Let us further suppose that he does regard the theory as doing a better job of conceptual unification of his past and present data than would be done by any rival theory of which he is aware. That still leaves him with two questions. **(1)** Are there things he could do—whether experimenting or merely waiting—which would bring him data in the light of which the outer-world theory would be inferior to some rival theory? **(2)** Are his past and present data handled less well by the outer-world theory than by some other theory which he just hasn't thought of?

Suppose now that there is a successful transcendental argument, A, which proves that self-knowledge requires the application of objectivity-concepts; and let us consider how far A can go towards answering those two questions.

In response to the first question, A would show that I could not find myself to be possessed of inner states which did not support an outer-world theory; and that is something which mere verificationism could not show. Still, it is not a very large gain: it does not even show that while I have self-knowledge I must inhabit this outer world, though of course a transcendental argument with a stronger conclusion than A's might show this.

Similar remarks apply to the second question, about superior rival theories which might be thought up now. If A is sound, then a self-conscious creature cannot rely on a non-outer-world theory in preference to any outer-world theory as a means for handling his inner states. Once more, however, A as it stands does not rule out the possibility that a self-knower should come to think that he has been quite wrong about what kind of outer world he inhabits.

It could be argued that A doesn't answer the second question at all: granted that a self-knower must manage his

inner states with the aid of an outer-world theory, perhaps he could at the same time acknowledge that some rival theory was superior? Isn't it conceivable that I should have to do my basic thinking in terms of theory T while realizing that theory T* was more powerful or more economical? I am not sure about this. If a self-knower must employ T, that must be because he needs conceptual services which T can render him; and so he cannot both need T and know a superior rival T* which renders all the main conceptual services which T renders. Objection: 'But he might know a rival T* which is over-all superior to T although it does not render all the services which T renders.' I cannot refute that, but it is so peculiar that I am prepared to walk away from it and to conclude that A does, near enough, show that we could not discover through intellectual effort that our past and present data are better handled by a non-outer-world theory.

Rorty assigns to transcendental arguments this role, as a kind of bulwark against rival conceptualizations for past and present data (Rorty, 1971, pp. 10–11). But he doesn't envisage their ever proving that a certain conceptualization of one's inner states is absolutely indispensable. At best, he says, one can hope to defend a favoured conceptualization against displacement by particular rivals, proving that C cannot be dislodged by C* because any use of C* would have to be parasitic on the use of C; and then for C** a fresh argument would be needed, and so on. I cannot refute this, but, as I said of a similar claim by Wilkerson, I see no reason to believe that transcendental arguments must operate in this way. Although I cannot actually produce anything fitting my specifications for 'argument A', I am not convinced that no such argument is possible.

11. Relevance to the foundations of science

Of the two answers which transcendental arguments can give to sceptical inquiries, the one about what I might find my future states to be like has no bearing on what concepts should be used in the foundations of science now. It is addressed not the question of what I can make, intellectually, of my present data, but rather to the question of what the future may bring. So it really belongs in the province of the so-called 'problem of induction': its only relevance to the foundations of science is that may help someone in whom Hume has caused a crisis of morale, leaving him unwilling to continue at all with an activity which may at any moment collapse. There may be sound transcendental arguments which set limits to how total and sudden that collapse could be known to be—that is, arguments proving that self-awareness is not compatible with experiential chaos or with rapid and radical changes in the patterns within one's experience. But the person who is demoralized by Hume is not typical of those who theorize about the conceptual foundations of science, and the nursemaid activity of consoling him is a rather lowly form of labour.

Incidentally, I don't think that any sound argument could rule out the possibility that my inner states might gradually change in such a way that although at any given time they were fairly efficiently handled by some scientific theory, the content of my *Weltanschauung* slowly changed until at last it had nothing in common with the one I accept now. So if someone is demoralized by the thought: 'Why bother with science if we cannot get results which are certain to remain at least approximately valid?', there is no consolation for him.

The second answer which transcendental arguments can give is *prima facie* more relevant to the conceptual foundations of science. If I cannot handle my data without the aid of concept C, not through the accidents of individual or cultural history but rather because C is conceptually required for self-knowledge, then there is no point in my looking for alternatives to C. It is a category, in roughly Kant's sense, and science is stuck with it.

If Rorty is right, one can never show that C has an absolutely categorical status, but only defend it against particular rivals. That would virtually put transcendental arguments out of business, for they would amount merely to the case-by-case evaluations which are what a foundations-of-science theorist engages in anyway.

Even if Rorty is wrong, I doubt if transcendental arguments should affect anyone's thinking about the conceptual foundations which science might have. This brings up the question of section 7 above, namely 'How much can be validly proved by a transcendental argument?' I cannot answer this, because I don't know just what is required for self-knowledge, or for any main species of it. I have argued only that self-knowledge including beliefs about the past requires the use of objectivity-concepts or something approximating to them; and although more than this can probably be shown, I don't know how much more. Still, I confidently guess that there is no concept C such that **(i)** C is provably required for self-knowledge or for some main species of it, *and* **(ii)** someone might seriously propose to give science new foundations which did not involve C. Kant tried to show that science must use the concepts of cause and substance in such a way as to be committed to strict determinism and to some conservation law; and those two theoretical

items might (and one of them did) come under challenge in a re-thinking of the foundations of science. But since Kant failed actually to prove that self-knowledge requires an acceptance of strict determinism or of a conservation law, this example does not refute my guess. Of course it is only a guess; but I am sure enough of it so that if I were actively working on questions about the conceptual foundations of science I would not look to transcendental arguments for help.¹

¹ For extremely helpful comments on an earlier draft of this paper I am indebted to Michael Beebe, Judith Jarvis Thomson, and Barry Stroud.

Bibliography

- A. J. Ayer, *The Concept of a Person*, London 1963.
- Jonathan Bennett, *Kant's Analytic*, Cambridge 1966.
- René Descartes, *Oeuvres philosophiques* (ed. by F. Alquié), Paris 1967, vol. II.
- Alan H. Goldman, 'Can A Priori Arguments Refute the Sceptic?', *Dialogue* 13 (1974).
- Moltke S. Gram, 'Transcendental Arguments', *Noûs* 5 (1971).
- A. Phillips Griffiths, 'Transcendental Arguments', *Proceedings of the Aristotelian Society*, suppl. vol. 43 (1969).
- Peter Hacker, 'Are Transcendental Arguments a Version of Verificationism?', *American Philosophical Quarterly* 9 (1972).
- Ross Harrison, *On What There Must Be*, Oxford 1974.
- Jaakko Hintikka, 'Transcendental Arguments: Genuine and Spurious', *Noûs* 6 (1972).
- David Hume, *Treatise of Human Nature* (ed. by L. A. Selby-Bigge), Oxford 1964.
- Immanuel Kant, *Critique of Pure Reason*.
- John Kekes, 'The Scandal of Philosophy', *International Philosophical Quarterly* 12 (1972).
- Stephan Körner, 'The Impossibility of Transcendental Deductions', *The Monist* 51 (1967).
- John Locke, *An Essay Concerning Human Understanding*.
- J. J. Macintosh, 'Transcendental Arguments', *Proceedings of the Aristotelian Society* 43 (1969).
- Norman Malcolm, 'The Conceivability of Mechanism', *Philosophical Review* 77 (1968).
- Richard Rorty, 'Strawson's Objectivity Argument', *The Review of Metaphysics* 24 (1970).
- Richard Rorty, 'Verificationism and Transcendental Arguments', *Noûs* 5 (1971).
- Jay Rosenberg, 'Transcendental Arguments Revisited', *The Journal of Philosophy* 72 (1975).
- Sydney S. Shoemaker, *Self Knowledge and Self-Identity*, Ithaca 1963.
- G. W. Smith, 'The Concepts of the Sceptic: Transcendental Arguments and Other Minds', *Philosophy* 49 (1974).
- William D. Stine, 'Transcendental Arguments', *Metaphilosophy* 3 (1972).
- P. F. Strawson, *Individuals*, London 1959.
- P. F. Strawson, *The Bounds of Sense*, London 1966.
- Barry Stroud, 'Transcendental Arguments', *Journal of Philosophy* 65 (1968).
- Judith Jarvis Thomson, 'Private Languages', *American Philosophical Quarterly* 1 (1964).
- Jeffrey Tlumak, 'Some Defects in Strawson's Anti-Skeptical Method', *Philosophical Studies* 28 (1975).
- T. E. Wilkerson, 'Transcendental Arguments', *Philosophical Quarterly* 20 (1970).
- Bernard Williams, 'Knowledge and Meaning in the Philosophy of Mind', *Philosophical Review* 77 (1968).
- Ludwig Wittgenstein, *Philosophical Investigations*, Oxford 1953.