CARTESIAN SKEPTICISM AND INFERENCE TO THE BEST EXPLANATION*

The problem of skepticism about the external world, or Cartesian skepticism, has its roots in the underdetermination of theory by evidence. We each adopt a body of common-sense beliefs about the world which answers to our sensory experience. In principle, however, the beliefs we base on that experience are subject to underdetermination, and we can devise radical alternatives to the common-sense account. Such alternatives take the form of skeptical hypotheses, like Descartes's fiction that his experiences are caused by an evil demon.

Certainly, when the choice arises, we hold to the common-sense view, and reject its skeptical competitors.[1] But what (epistemic) reasons can we have for doing so? In cases of underdetermination generally, principles of inference to the best explanation can license the choice of one theory over others. Accordingly, we would be justified in preferring the common-sense account to skeptical hypotheses, if the common-sense account provides better explanations of why our experience is the way it is.[2] My purpose here is to inquire into the explanatory advantages of the common-sense view, and to develop a response to skepticism along the lines just indicated.[3]

One obstacle to carrying out this project is that the standards by

* To be presented in an APA symposium on Cartesian Skepticism, December 30. Berent Enc will comment; see this JOURNAL, this issue, 667–8. I would like to thank Phillip Bricker, Anthony Brueckner, Jeremy Butterfield, Stewart Cohen, Richard Feldman, Mark Johnston, Peter Lipton, Thomas Tymoczko, and Daniel Velleman. Research for this paper was supported by the American Council of Learned Societies and the National Endowment for the Humanities.

[1] Some philosophers, especially followers of Ludwig Wittgenstein, would deny that skeptical hypotheses can genuinely compete for acceptance with the body of our common-sense beliefs. See, for example, Stanley Cavell, *The Claim of Reason* (New York: Oxford, 1979), pp. 218–220.

[2] This approach to skepticism has been advocated by Michael Slote, Frank Jackson, Jonathan Bennett, James Cornman, J. L. Mackie, and Alan Goldman, among others.

[3] On certain views about skepticism and about inference to the best explanation, this approach to skepticism will seem ill-conceived. One might hold that it is simply constitutive of rationality to reject skeptical hypotheses out of hand; thus, it is unnecessary to enter into the relative explanatory merits of the common-sense view and its skeptical alternatives. From another point of view, the explanatory advantages of the common-sense view could never give us a reason to accept it as *true,* rather than as merely handy or to our taste. The issues that arise here are important, and they must be addressed at some point by anyone who bases an answer to skepticism on explanatory considerations. These very general objections will be moot, however, if the appeal to explanatory considerations does not even succeed on its own terms. Whether it does so is my present concern.

0022-362X/90/8711/658–666

which explanations are evaluated are themselves difficult to identify and to make precise. In what follows, I shall be making some controversial assumptions about explanatory goodness, and I shall have to rely on largely unanalyzed notions of simplicity, ad-hoc-ness, and the like. To be explicit, I shall presuppose:

(a) Ad hoc explanations should be avoided, i.e., very roughly, if *A* is offered as an explanation of *B*, *A* ought not to be isolated from other explanations and data (it ought to be independently testable, it must figure in the explanation of something other than *B*, etc.).

(b) Other things being equal, a simpler explanation is superior to a more complicated one.

(c) Where explanation is concerned, more is better, if you get something for it. In particular, it is desirable to be able to give higher-level explanations of lower-level ones.

Another methodological point requires some comment. In comparing skeptical hypotheses with our everyday account of the world, I shall exclude from the latter any advanced scientific beliefs. To be sure, science adds great power and coherence to our explanations of phenomena, and one might argue that no explanatory scheme the skeptic devises could seriously compete with our best scientific theories. But it seems implausible that, without such theories, we would lack adequate grounds for rejecting skeptical hypotheses. Accordingly, I shall try to show that even a scientifically unsophisticated common-sense view of the world provides more adequate explanations than its skeptical competitors.

I

Our beliefs about the external world serve an explanatory function. A person's sensory experience exhibits patterns and regularities at many levels, and our common-sense beliefs account for these in ways that seem to be coherent and economical. I shall call the body of these beliefs the *real-world hypothesis* (RWH).[4]

The skeptic points out that there are alternative explanations of how a person's sensory experience arises. In principle, a great many ways of formulating and developing these counterhypotheses are open to the skeptic—for example, through various stories about evil

[4] This way of putting things may seem unfortunate to those who reject the representative theory of perception. But the point could be recast as follows: we have a set of beliefs about the world, i.e., the RWH. Our having those beliefs admits of alternative explanations, including skeptical explanations and the RWH itself. The tenability of skepticism turns on whether the truth of the RWH provides a better explanation than do skeptical hypotheses of why we believe the RWH in the first place.

demons and brains in vats. But elaborate (not to mention crazy) fantasies of deception may be only tenuously connected to the content of one's experiences and may lack cohesiveness. For instance, suppose that you seem to see some snow falling, and the skeptic suggests that this experience is being foisted on you by a demon. Then, to explain why the demon makes you have snow experience (rather than experience of some other kind), the skeptic tells you that there is a second demon that has put the first one up to it. Clearly, we are not getting anywhere; positing a second demon that directs the first on this occasion (and does only that) is explanatorily idle or ad hoc. The skeptic could try to escape such a result by refusing to say in any detail how your experiences come about. A hypothesis in this vein might specify only that your experiences are all caused by some deceptive spirit, and no more. The cure is as bad as the disease, however: the skeptic will succeed in avoiding ad hoc higher-level posits only by foregoing higher-level explanations altogether.

The RWH, by contrast, gives us a rich and well-integrated explanatory apparatus. We not only posit objects that cause our experiences, we are also able to explain why and how these objects behave as they do. If the explanations provided by a skeptical counterhypothesis are either ad hoc or impoverished in comparison with those of the RWH, then we have good grounds for preferring the latter to the former. According to the skeptic, we fail to know things because the RWH is faced with competitors that we have no reason to reject. But we have just seen that not any competitor will do. The skeptic's position will be empty unless he can provide us with reason to think that a *satisfactory* competitor exists (in particular, a sufficiently rich competitor that is not unduly burdened with ad hoc explanatory posits).

The lesson here is that the skeptic needs to frame an alternative that matches the RWH very closely. If a skeptical hypothesis can be made sufficiently similar in relevant respects to the RWH, then, one might expect, that skeptical hypothesis will match the RWH in explanatory adequacy. To the extent that explanatory virtues like coherence, depth, and simplicity are matters of theoretical structure, a skeptical hypothesis that is isomorphic to the RWH will explain things just as well as the RWH does. An improved skeptical hypothesis of this sort has to satisfy two principal constraints: (i) it should invoke items corresponding to the elements of the RWH; (ii) it should also posit, as holding of these items, a pattern of properties, relations, and explanatory generalizations mirroring those of the RWH.

As an example of how this would work, suppose you seem to see the wind blowing a piece of paper off your desk. According to the

RWH, your visual impressions of the paper flying off your desk are caused by the paper. Similarly, your tactile sensations of the wind are caused by a real movement of air against your skin. And, finally, the wind stands in a relation of cause and effect to the movement of the paper. The skeptic's procedure will be to extract the explanatory skeleton or core from the RWH—that there are *some* entities bearing *some* properties that are related in ways exactly analogous to those specified by the RWH—and then to add that the entities and their properties are somehow different from the ones mentioned in the RWH.[5]

Thus, a skeptical hypothesis might present the following alternative explanation of your experiences. All that there is to the world is your brain in a vat, and a computer that is connected to your brain. Your tactile experiences are caused by the realization of a computer program that simulates wind, and your visual impressions are caused by the realization of another program that simulates a paper blowing off a desk. Also, the skeptical hypothesis can specify that the first routine calls the second, so that (as in the RWH) the cause of the wind experience would be the cause of the cause of the paper-blowing experience. This way of reconstructing the explanatory structure of a small fragment of the RWH might be extended to apply to all the entities and explanatory connections posited by the RWH. The result would be a skeptical hypothesis that was completely isomorphic to the RWH, with portions of the computer disk supposed to occupy the explanatory roles we normally assign to familiar objects.[6] I shall be calling this the *computer skeptical hypothesis* (CSH).

Of course, the CSH is an outlandish suggestion, and we are confident that it is false. Yet, in reflecting on this situation philosophically, it is possible to misread what has gone wrong. One proposal is that skeptical hypotheses are invariably burdened with more unexplained

[5] Basically, this amounts to something like forming the Ramsey sentence of the RWH and adding to it further specifications that, in each case, the object or property denoted by the bound variables is something other than the one posited by the RWH. The RWH itself can be construed as the "Ramsey sentence" plus the stipulation that the objects and properties called for by the "Ramsey sentence" are indeed the familiar ones. See here Grover Maxwell, "Theories, Perception, and Structural Realism," in R. Colodny, ed., *The Nature and Function of Scientific Thought* (Pittsburgh: University Press, 1970) and for some needed refinements, David Lewis, "How to Define Theoretical Terms," *Philosophical Papers,* vol. I. (New York: Oxford, 1983). The possibility of framing skeptical hypotheses with the same structure as the RWH is noted by Lawrence Sklar in his "Saving the Noumena," *Philosophy and Spacetime Physics* (Berkeley: California UP, 1985), pp. 59–60.

[6] We need not suppose that the computer itself was built or programed by anyone. Rather, this hypothesis is to be understood simply as a description of an alternative way the (physical) world might be.

explainers than is the RWH.[7] The CSH will lack answers to questions like 'Why does the computer operate the way it does?' or 'Where did the computer come from in the first place?' But it is not at all clear that the RWH does any better in the face of analogous demands. Both the CSH and the RWH invoke ultimate regularities that are not themselves explained, and neither can account for the existence of the physical world as such. Generally, since the RWH and the CSH are meant to have the same structure, anywhere the RWH can explain a lower-level phenomenon by a higher-level regularity, the CSH should be able to do the same. The CSH will have unexplained explainers only insofar as the RWH has them also.

Another suggestion that enjoys some currency is that the RWH is, in a very straightforward way, simpler than the CSH, and hence to be preferred. The idea here is that there would be a one-one mapping from the objects posited by the RWH to their stand-ins in the computer's memory, where these are treated as discrete individuals. There are, though, items required (at least tacitly) by the CSH which escape this mapping, e.g., the computer's central processing unit and perhaps the brain in the vat itself. So, the argument runs, the CSH is committed to the existence of more items than the RWH, and is to be preferred on that account.

This line of thought is problematic in several respects. First, one could just as well argue that the CSH is simpler than the RWH, on the grounds that the CSH posits only two objects (the computer and one's brain), whereas the RWH is committed to the existence of a great many more things. Moreover, it is far from clear that, all by itself, positing fewer entities is a theoretical virtue.[8] And finally, if need be, the CSH could be revised to eliminate the role of the central processing unit altogether. The skeptic could suppose that the elements of the computer memory act directly on each other, and on the seat of consciousness, in causal patterns that mirror those of the RWH.

Now, as will emerge shortly, I think there is something right about the claims that the CSH is less coherent and less simple than the RWH. But if explanatory coherence and simplicity are treated solely in structural terms, it should not be surprising that these claims do

[7] A claim of this sort is made by Alan Goldman, although it is directed at a fantastical skeptical story that postulates experimenters with deceptive motives. See Goldman, *Empirical Knowledge* (Berkeley: California UP, 1989), p. 212.

[8] It could be objected that what matters for explanatory adequacy is not economy with respect to the number of individuals posited, but rather in the number of different kinds invoked. But this does not appear to help—the skeptic can get by with just a few kinds of things (brain, vat, computer) while the RWH might be said to invoke these and many more.

not go through. After all, the causal-explanatory structures invoked by the RWH and the CSH are identical; the two differ only as to what entities bear the specified causal relations to one another.

The rejoinders just considered miss something important about the motivations behind the skeptic's argument. At root, the skeptic questions our ability to read off the "real" or intrinsic character of things from those things' causal behavior. This challenge emerges in its simplest form with the initial thought that one's experience of any familiar object might be caused by something other than that object (e.g., an evil demon). The point is that the known effect—namely, your experience—does not fix the character of its cause.[9]

On the face of it, the requirement that a skeptical hypothesis must have a more fully articulated structure—one that matches the RWH in various ways—seems insufficient to meet this problem. For, if we assume that causal relations are contingent and that there is in principle no obstacle to our positing whatever causal relations we like, what reason could there be why one set of entities is better suited than another to occupy the positions within the structure of the RWH itself? It would appear that, in principle, there should be skeptical hypotheses that will explain the contents of one's experience just as well as the RWH. The choice between such hypotheses and the RWH will then be arbitrary, giving the skeptic what he needs.

II

To appreciate the superiority of the RWH over its skeptical competitors, we need to take into account the content, as well as the form, of the explanations the RWH provides. In particular, our ordinary view of things involves beliefs in the existence of objects with familiar spatial characteristics (e.g., we believe that there are bricks that are oblong and oranges that are round). The ascription of specific spatial properties to objects does explanatory work within the RWH (e.g., accounting for why oranges roll easily and bricks do not). Since the CSH posits objects with altogether different spatial characteristics —we are assuming that its objects are just portions of a computer disk—the CSH will have to account for the relevant phenomena in some other fashion. But by bringing in these additional explanations (whatever they may be), the CSH runs the risk of taking on a more elaborate explanatory apparatus than the RWH. To put the point I am trying to make more directly: niceties aside, the fact that some-

[9] This was the way Kant understood the situation. The skeptic, he says, "assumed that the only immediate experience is inner experience and that from it we can only *infer* outer things—and this, moreover, only in an untrustworthy manner, as in all cases where we are inferring from given effects to determinate causes"; *The Critique of Pure Reason,* N. K. Smith, trans. (New York: St. Martin's, 1965), p. 245.

thing is spherical explains why it behaves like a sphere (in its interactions with us and with other things). If something that *is not* a sphere behaves like one, this will call for a more extended explanation.

This intuitive claim is bound to raise some philosophical qualms. Why must the fact that the CSH invokes *different* configurations of matter in its explanations mean that CSH has to be *more complicated* than the RWH? Again, setting niceties aside, why is the skeptic not free to stipulate that, in his account, it is certain magnetic patterns on a disk, not spheres, which behave like spheres (at least in terms of the experiences they bring about, directly and indirectly)?

Let us see just what would be involved in maintaining an explanatory parity between the CSH and RWH. To fix ideas, suppose that, according to the RWH, there is a hyacinth beside your doorway. For each RWH object, there has to be a CSH counterpart, which we can imagine to be the piece of the computer disk which stores the information about the object to be simulated. So, the CSH would have it that there is a piece of the disk holding a file about a hyacinth beside your door, specifically. Moreover, wherever the RWH assigns a certain property to the hyacinth, the CSH must ascribe a corresponding, but different property to the hyacinth's CSH analog. According to the RWH, the hyacinth has a particular location, namely, that of being beside your door. The hyacinth counterpart will have some parallel feature, which we might call a "pseudo location." The pseudo location of the hyacinth counterpart is just that physical property in virtue of which the counterpart simulates being located near your door. In general, what the RWH explains by reference to genuine locations, the CSH will explain in terms of these pseudo locations.

Since we make reference to the locations of objects in giving various everyday explanations, location properties are part of the explanatory apparatus of the RWH. Now, we find that the (genuine) locations ascribed to any two objects at a time are invariably different. We do not need any empirical law or regularity to explain this; it is a necessary truth pertaining to the nature of physical objects that there cannot be two such objects at the same place at the same time.[10]

The explanatory structure of the CSH is meant to duplicate that of the RWH. Since the CSH is isomorphic to the RWH, and the RWH always ascribes different locations to the objects it posits, the CSH will invariably ascribe different pseudo locations to things it posits.

[10] For a discussion of this principle, see Denis Robinson, "Re-identifying Matter," *The Philosophical Review*, XCI (1982): 317–341; on the role of necessary truths in explanations, see Clark Glymour, "Explanation and Realism," in J. Leplin, ed., *Scientific Realism* (Berkeley: California UP, 1984), esp. pp. 184–6.

This calls for an explanation, if possible. At this point, however, the CSH faces a loss in either simplicity or explanatory power. To make the issue more concrete, imagine that the way things work in the CSH computer is that each object's pseudo location is the physical realization of having coordinates (x,y,z) written in its file.[11] There will have to be some explicit principle within the CSH that no two objects are to be assigned the same pseudo location, i.e., that no two objects are to have the same coordinates written in their files. Otherwise, the fact that no two objects have the same pseudo location remains unexplained. Of course, the CSH would include within it the necessary truth that two physical objects cannot occupy the same *genuine* location in space, but this is of no help to the CSH in explaining why two of its objects cannot have the same *pseudo* location. To achieve this, it would appear that the CSH has to add an extra empirical regularity, to which no regularity in the RWH corresponds. Such an addition will make the CSH inferior to the RWH on simplicity grounds, however.

The skeptic could escape this outcome if it could be guaranteed by some other necessary truth that different CSH objects will have different pseudo locations. In other words, the pseudo location of a CSH object would have to be encoded by some physical property P (other than that of having some specified location), such that it is impossible for two physical objects to have P at the same time. But it seems to me that there are no such physical properties. After all, if a given physical object with whatever properties exists at one place, it appears perfectly possible for there to be an absolutely similar object elsewhere, instantiating all the same properties at the same time—except location.[12]

Actually, the problem facing the skeptic is a general one, independent of the fact that CSH itself invokes physical objects (i.e., bits of computer disk) in its explanations. Suppose that the skeptic offers instead a quasi-Leibnizian hypothesis, according to which the world consists solely of minds and their properties. These minds and their states are supposed to act in ways that mirror the behavior of everyday things as specified by the RWH. Each mind that stands in for a RWH object must have a property corresponding to the genuine location the RWH ascribes to its object; this pseudo location will be a (partial) mental state. The question arises again as to why these

[11] For purposes of exposition, I am pretending that an object is located at a point rather than a region.

[12] Of course, there are characterizations like 'the only building taller than 110 stories' or 'identical to Socrates', which are satisfied by at most one object at a time. If these expressions involve reference to properties, they are properties of a different type than those with which I am concerned here.

pseudo locations are invariably different from one mind to another. Presumably (*pace* Leibniz), it is possible for two different minds to think exactly the same thing at the same time, so no necessary truth prevents them from having the same pseudo location. Once again, such an occurrence would have to be ruled out by some kind of extra "exclusion principle," for which no counterpart exists in the RWH.

I claimed earlier that our normal ascription of spatial properties to things does real explanatory work; furthermore, it seems plausible that you incur an added explanatory burden if you suppose that something lacking a particular spatial property still behaves as though it had it. What I have been saying about locations and pseudo locations makes this same point on a more abstract level. In skeptical hypotheses, some other property (e.g., a magnetic property or a mental property) is supposed to substitute for the spatial property of being located at particular place. As we have seen, further explanation is then needed to establish why these properties, which are not genuine location properties, behave as though they were. It seems that this sort of difficulty will attach to skeptical hypotheses generally, giving us good reasons to reject them.[13]

III

I have presented some antiskeptical arguments based on explanatory considerations. But surely there is a world of familiar objects about us, and we have known that all along. So what, then, is the point of giving these arguments in the first place? This question deserves an extended response, but for now a very brief answer will have to do. I take it that the specious character of the explanations the skeptic offers is immediately apparent—they come across as contrived or unduly indirect—and this is a reason why we reject skepticism as a doctrine. Realizing that skeptical hypotheses are defective, however, is not the same thing as spelling out precisely what their defects are. To do this requires philosophical work—work of the sort I have undertaken here.

JONATHAN VOGEL

Amherst College

[13] One might try to frame a skeptical hypothesis that avoids this difficulty by assigning to objects different locations (and spatial properties generally) in place of those specified by the RWH. Formulated this way, our problem becomes one of choosing a particular geometry of the world from among those logically compatible with the empirical data, and one might continue to defend the choice of the RWH by appeal to explanatory considerations. See here Lawrence Sklar, *Space, Time, and Spacetime* (Berkeley: California UP, 1977), pp. 91–101, although Sklar himself is highly critical of such uses of inference to the best explanation. Sklar has a valuable discussion of the affinities between Cartesian skepticism and problems in the epistemology of geometry.