PETER VAN INWAGEN

LEHRER ON DETERMINISM, FREE WILL, AND EVIDENCE

(Received 2 August, 1971)

Keith Lehrer has offered a novel and ingenious epistemological argument for the compatibility of free will and determinism.[1] David Blumenfeld has recently claimed that this argument begs the question.[2] I shall argue that, while Blumenfeld's claim is mistaken, Lehrer's reasoning is nonetheless unsatisfactory.

Lehrer argues that we could collect evidence on which it would be highly probable that an agent could have performed some act he did not perform. He asks us to imagine that we carry out an experiment in which a certain person (I shall call him 'S') is required to lift his arm many times and under a great variety of circumstances.[3] As the hours pass, we compile a record of our observations of S's arm-liftings; we enter into the record hundreds of observed arm-liftings, and the hundreds of circumstances under which they were performed. We then tell S to lift his arm twice in rapid succession. We see him lift his arm once, and then we briefly close our eyes, so that we do not see whether he lifts it a second time. (Let T be the time at which our eyes were closed.) We then examine recording instruments and see that S's condition at T (blood pressure, alpha rhythm and so on) was very similar to what it was when we saw him raise his hand just before T. Let E be all the evidence we have collected in our experiment up to this point. Lehrer argues that the hypothesis H that S was able to lift his arm at T is *highly probable* on E. In order to show this, he proposes certain conditions that evidence like E must satisfy if it is to render hypotheses like H highly probable. He claims that these conditions are 'typical of the usual canons of inductive evidence', and that, in fact, given the usual canons of inductive evidence, E is excellent evidence for H.[4] Rather than examine the specific epistemic conditions that Lehrer proposes,[5] I shall offer a brief defense of this last, more general, claim. Lehrer does not argue explicitly for the conclusion that the usual canons of induction support his thesis

Philosophical Studies **23** (1972) 351–357.

that H is highly probable on E, but such an argument is not hard to construct.

Surely the following is a reasonable inductive principle: The hypothesis that x has property p at time t is highly probable on evidence consisting in a set of observations entailing (i) that x had p at many times other than t, (ii) that x had p under a great variety of conditions, (iii) that x had p a short time before t, and (iv) that x changed very little in other respects between t and the latest time before t at which x was observed to have p, and *not* entailing that there was any time at which x did not have p. Now suppose that x is S and t is T and p is the property of being able to lift one's arm. If our inductive principle is correct, then H is highly probable on E,[6] since on each occasion on which we observed S lifting his arm, our observation entails that on that occasion S had the ability to lift his arm. Moreover, there is no need actually to carry out this experiment, since it is obvious without putting it to the test that we could collect just such evidence as E if we wished.

Now suppose that in fact S *did not* raise his arm at T, and, moreover, says that he did not even *try* to raise his arm at T. Let E′ be the conjunction of E with the statement that S did not raise his arm at T and says that he did not try to raise his arm at T. Is H highly probable on E′? It should seem so, since if we observe that a person does not perform some act A, and if he tells us that he did not *try* to perform A, then this observation is, when taken by itself, *irrelevant* to deciding whether he could have performed A; it supports neither the hypothesis that he could have performed A nor its denial. And if H is highly probable on E′, then H′, where H′ is the hypothesis that S could have raised his arm at T and did not raise his arm at T, is highly probable on E′. I shall call Lehrer's argument for this conclusion 'Argument A', and the remainder of his argument 'Argument B'. I shall formulate Argument B as a sequence of numbered statements, since I shall wish to obtain related arguments from it by various simple logical manipulations:

B(1) H′ is highly probable on E′.
B(2) Indeterminism is not highly probable on E′.
B(3) If p is highly probable on evidence e, and q is not highly probable on e, then p does not entail q.
$\therefore$B(4) H′ does not entail indeterminism.

It is obvious that if free will were inconsistent with determinism (or, as we shall say, if *incompatibilism* were true), then B(4) would be false, since if a statement asserting that S was able to do what he did not do is inconsistent with determinism, then it entails indeterminism. Therefore, Lehrer's argument has the consequence that incompatibilism is false. Premise B(1) was established by Argument A. Premise B(3) is a corollary of a theorem of the probability calculus. And B(2) is obviously true, Lehrer argues, since E′ is utterly irrelevant to indeterminism.

Now an incompatibilist is not likely to find Lehrer's reasoning convincing. He will almost certainly object to B(1), and, if pressed to support his objection in the face of Argument A, might simply stand Argument B on its head to produce what I shall call Argument C, having premises C(1) (= ~B(4)), C(2) (=B(2)), C(3) (=B(3)), and conclusion C(4) (= ~B(1)). And this would not be simply to beg the question, since, presumably, the incompatibilist has some independent argument for C(1).

Can Lehrer reasonably be accused of begging the question against the incompatibilist? Blumenfeld claims that B(1) "assumes the very point at issue", and is really no more than the articulation of a popular prejudice.[7] (That is, he makes this claim about the corresponding premise in his formulation of Lehrer's argument.) For suppose, he argues, that determinism and incompatibilism are both true: "nothing in this state of affairs would rule out, or even make improbable that we should be able to run Lehrer's experiment just as he imagined it".[8] And having run Lehrer's experiment we could go on, using Lehrer's argument, to 'prove' compatibilism. But this hardly shows that Lehrer's argument begs the question. Let us consider an analogy. Suppose I argue that the Taj Mahal is white even when no one is observing it, in opposition to some philosopher, P., who holds that physical objects have colors only while they are being observed. I imagine an experiment in which a great many observations are made of the Taj Mahal's color, and argue that, given the usual canons of induction, it would be highly probable on such evidence that the Taj Mahal was white at T′ when no one was looking at it. Am I begging the question against P.? This question, like most questions about whether the question is being begged, is not very clear. But one thing *is* clear: one cannot show that I am begging the question simply by pointing out that, even if P.'s thesis were true, I should still be able

to run my experiment and obtain exactly the same results. This is only to point out that induction might fail, and this, everyone agrees, is logically possible. And just as induction would fail with respect to color in a world in which the Taj Mahal is white only when its color is being observed, induction, according to Lehrer, would fail with respect to abilities in a world in which S is able to lift his arm only when he does lift his arm. Now the incompatibilist might wish to reject the universal applicability of principles of induction that Lehrer accepts as universally applicable, and, since no inductive principle is a necessary truth, he may do so without contradiction. Nevertheless, such a rejection is not to be undertaken lightly; he cannot dismiss these principles as mere prejudices of the vulgar. Moreover, Lehrer will almost certainly find at least one of the incompatibilist's premises less plausible than his own inductive principles.

In fact, Lehrer and the incompatibilist are equally vulnerable to charges of question-begging, and for similar reasons. If the incompatibilist can accuse Lehrer of begging the question against incompatibilism by assuming the universal applicability of the principles of induction, then Lehrer can accuse the incompatibilist of begging the question against compatibilism by assuming the truth of whichever of his premises it is that Lehrer happens to dislike most. But this would be pointless. The real question is not whether either philosopher has begged the question, but (assuming that the premises common to arguments B and C are true) whether Argument A or the incompatibilist's argument for C(1) is the *better* argument.

It is not altogether clear, however, that both the premises common to arguments B and C *are* true. I shall argue that *either* Argument A is unsound and B(1) is false, or if B(1) is true, then the truth of B(2) is not so obvious as Lehrer imagines. In order to show this, I shall first construct an argument for the denial of B(1). This argument will be a modification of Argument C, consisting essentially in the replacement in Argument C of indeterminism by a thesis obviously entailed by H′.

Let *demonism* be the thesis that there exists a demon who can infallibly predict the future and who behaves as follows. Whenever any person A is born, this demon makes a list of every overt act that A will ever perform. He gives this list to an obedient lesser demon, and instructs him to keep a close watch on A. As long as A behaves in accordance with the list, the

subordinate demon is to keep himself concealed from all human observation. But if A should try to perform some act not on the list, the demon must prevent him from succeeding by main force (and he is able to do this). Let us again consider our hypothesis H′. It is clear that if H′ is true, then demonism is false, since if demonism were true, H′ would be false. For if S did not raise his hand at T, then this would have been predicted, and the subordinate demon would, in effect, have been instructed to *insure* that S does not raise his hand at T. We now argue (Argument C′):

C′(1) H′ entails the denial of demonism.
C′(2) The denial of demonism is not highly probable on E′.
C′(3) (=C(3)=B(3))
∴C′(4) H′ is not highly probable on E′. (=~B(1))

We may justify C′(2) just as Lehrer justified B(2): evidence E′ is utterly irrelevant to the question whether demonism is true. Therefore, it should seem, we must reject B(1), since Argument A, though perhaps plausible, is surely less convincing than Argument C′.

Lehrer, however, has suggested to me that the strategy of Argument C′ must be faulty, since it can be used to show that no hypothesis is highly probable on any evidence. Suppose, for example, that a certain urn initially contained one thousand balls, each one either black or white, and that I have drawn from it nine hundred ninety-nine white balls. Let evidence E″ comprise just this information. Let W be the hypothesis that the ball remaining in the urn is white. Surely if any hypothesis is ever highly probable on any evidence, then W is highly probable on E″. But let B be the hypothesis that whenever there is one and only one ball in an urn, a demon turns it black (if it is not already black). We argue (Argument D):

D(1) W entails the denial of B.
D(2) The denial of B is not highly probable on E″.
D(3) (=B(3))
∴D(4) W is not highly probable on E″.

Now D(4) is surely unacceptable, and the only premise of Argument D that it is possible to reject is D(2). We must therefore conclude that the denial of B is highly probable on E″, despite the initial implausibility

of this conclusion. But in that case, it is by no means evident that our argument for C′(2) is sound. This conclusion, however, is at best a Pyrrhic victory for Lehrer. For, it should seem, either B(1) is false, in which case Lehrer's argument is unsound, or, if B(1) is true, then a very strong argument will be needed to show that indeterminism is not highly probable on E′. Mere subjective plausibility is no more sufficient to support B(2) than it is sufficient to support C′(2) and D(2). But subjective plausibility is all that Lehrer offers:

> ... However, it is incongruous to suppose that the *kind* of evidence we would obtain in our imaginary experiment... would disprove the thesis of determinism... this evidence hardly seems sufficient to justify the hypothesis that [S's] behavior was not causally determined![9]

In this passage, Lehrer confuses the issue by talking of proof and justification. The question is not whether E′ *justifies* or *proves* indeterminism (as indeed it does not), but whether indeterminism is *highly probable* on E′. An hypothesis may be highly probable on evidence that fails to justify it, simply because the evidence is irrelevant. For example, while we must grant that the denial of B is highly probable on E″, we should hardly want to say that E″ *justifies* the denial of B.

Lehrer's argument is simply not strong enough to show that indeterminism is not highly probable on E′. For indeterminism *is* highly probable on E′ (given that B(1) is true) just in case H′ entails indeterminism. And this truth cannot be used to show by *Reductio ad Absurdum* that H′ does not entail indeterminism, since, as we have seen, the statement that indeterminism is highly probable on E′ is absurd only on the mistaken supposition that an hypothesis cannot be highly probable on irrelevant evidence.

Lehrer's argument, I think, shares one feature of the traditional arguments for the existence of God: though it may strengthen the faith of the believer, it will leave even the most rational infidel unconverted.

Syracuse University

NOTES

[1] 'An Empirical Disproof of Determinism?' in *Freedom and Determinism*, (ed. by Keith Lehrer), (Random House, New York, 1966), pp. 175–202.

[2] 'Lehrer's Proof of the Consistency Thesis', *Philosophical Studies* **22** (1971),. 26–30.

[3] I have changed various details of Lehrer's imagined experiment, and of his argument. The symbols (S, E, H, etc.) are not Lehrer's.

[4] *Op. cit.*, p. 181.

[5] These are: *temporal propinquity*, *circumstantial variety*, *agent similarity*, and *simple frequency*. See Lehrer, *op. cit.*, pp. 178–80, or Blumenfeld, *op. cit.*, p. 27.

[6] Since Lehrer assumes that "if a hypothesis is very highly probable with respect to some kind of empirical evidence, then it is possible to know that hypothesis empirically," (*Op. cit.*, p. 117, *n* 3), he makes the claim that we can *know* H empirically on the basis of E. We shall see later that this assumption is unjustified. The stronger claim, however, is not necessary for his argument.

[7] *Op. cit.*, p. 29. I am not quite sure what 'the point at issue' is supposed to be. It cannot be compatibilism, since, in that case, B(1) alone would entail compatibilism. But B(1) entails compatibilism only in conjunction with two other premises for which Lehrer has arguments that do not depend on B(1).

[8] *Loc. cit.*

[9] *Op. cit.*, pp. 198–9.