INDUCTION AND REASONING TO THE BEST EXPLANATION*

R. A. FUMERTON

University of Iowa

In this paper I want to cast doubt on the claim that there is a legitimate process of reasoning to the best explanation which can serve as an alternative to either straightforward inductive reasoning or a combination of inductive and deductive reasoning. I shall argue a) that paradigmatic cases of acceptable arguments to the best explanation must be considered enthymemes and b) that when the suppressed premises are made explicit we have all of the premises we need to present either a straightforward inductive argument or an argument employing both induction and deduction.

In attempting to justify his belief in the existence of certain sorts of entities the philosopher sometimes appeals to an argument to the best explanation. Thus, for example, Locke (1959, Book IV, Chap. XI, esp. pp. 328–329) and Russell (1959, pp. 22–23) both argue that one's justification for believing in the existence of objects distinct from, but somehow revealed to us through, fleeting and subjective experience is the fact that the former provide us with the best explanation for the occurrence of the latter.[1] Those who begin with the external world and worry about the physicist's justification for believing in the existence of such exotic and unobservable entities as electrons, quarks, gravitational fields and the like, again sometimes find themselves appealing to the existence of such things as the best explanation for various macroscopic phenomena. What is interesting and important about this kind of reasoning is that it is alleged to provide us with an alternative to inductive reasoning.[2] Indeed it is

*Received December 1979.

[1]Locke does not say in so many words that he is employing an argument to the best explanation, but I think this is the most reasonable explication of what he is doing. For a more recent attempt to offer the physical world as the best explanation for the occurrence of sensations, see (Mackie 1969).

[2]When I refer to inductive reasoning here and throughout this paper I am referring to reasoning roughly in accordance with the principles of induction set out by Russell in (1959, pp. 66 and 67). I am, of course, aware that most philosophers take such principles as far too simplistic to serve as the canons of inductive reasoning. The most serious objections center around the allegations that they allow obviously bad arguments. If the focus of this paper were inductive reasoning I would be prepared to argue that if we keep in mind the fact that the conclusion of a good inductive argument is probable only *relative to* its premises one can do a great deal to diffuse

Philosophy of Science, 47 (1980) pp. 589–600.

offered as an alternative which can give us justification where inductive reasoning cannot, thereby dissolving certain philosophical problems which puzzled those with too narrow a conception of what constitutes legitimate reasoning.[3]

In this paper I want to cast doubt on the claim that there is a legitimate process of reasoning to the best explanation which can serve as an alternative to either straightforward inductive reasoning or a combination of inductive and deductive reasoning.[4] The paper will be divided into three parts. In Part I, I shall argue a) that commonplace, paradigmatic cases of acceptable arguments to the best explanation must be considered enthymemes and b) that when the suppressed premises are made explicit we have all of the premises we need to present either a straightforward inductive argument or an argument employing both induction and deduction. In Part II, I shall consider the claim that the structure of reasoning to the best explanation is more complex than my treatment of it in Part I would suggest, and shall examine one specific attempt to flesh out an alternative model. Finally, in Part III, I shall briefly comment on what I suspect is the main concern of those who seek to find in reasoning to the best explanation an alternative to inductive reasoning.

I

When one presents an argument to the best explanation in ordinary discourse it often has the following form:[5]

---

the most obvious counterexamples. In any event I do not think any of the points I wish to raise concerning the distinction between inductive reasoning and reasoning to the best explanation will be seriously affected if we understand inductive reasoning roughly the way Russell did. I am also, of course, not forgetting that there is a problem about justifying *inductive reasoning*. In the framework of this paper I am assuming that there is a solution to that problem, perhaps the solution Russell, himself, suggested (1959).

[3]Thus those who hold that reasoning to the best explanation constitutes an alternative to inductive reasoning could cheerfully acknowledge Hume's claim that "we may observe a conjunction or a relation of cause and effect between different perceptions, but can never observe it between perceptions and objects" (1888, p. 212) but deny that this means one can only reason from the existence of certain perceptions to the future occurrence of still other perceptions. Similarly the scientific realist could admit that one can never establish what one can observe as evidence for what one cannot observe inductively, but deny that this forces him to choose between skepticism and some form of instrumentalism.

[4]Gilbert Harman (1965) agrees that inductive reasoning and reasoning to the best explanation collapse but tries to reduce inductive reasoning to reasoning to the best explanation. If the arguments that follow are sound the distinction between inductive reasoning and reasoning to the best explanation collapses in the *opposite* direction. For less radical criticisms of Harman's position and a reply by Harman see (Ennis 1968) and (Harman 1968).

[5]This is the argument form that Peirce called hypothesis or abduction and which he also sometimes referred to as reasoning to the best explanation. See (Peirce, 1931,

(A) 1) Q is the case
2) If P were the case Q would be the case [6]
3) P is the case

To consider a very simple example we might, upon coming across some footprints on a beach, reason to the conclusion that a man walked the beach recently employing the following argument:

(I) 1) There are footprints on the beach.
2) If a man walked the beach recently there would be such footprints.
3) A man walked the beach recently.

Though the argument hardly has the dramatic character of those with which Holmes used to amaze Watson, nevertheless we would presumably all take the inference from 1) and 2) to 3) to be a legitimate, if pedantic, piece of reasoning.

Now it goes without saying, I suppose, that if the inference from premises to conclusion in arguments having the above form is legitimate, the conditional employed in the premises must be other than material implication. Arguments having the form:

1) Q
2) $P \supset Q$
3) P

are obviously bad arguments even if the relation between premises and conclusion is only supposed to be one of confirmation. The inference from 1) If there is a unicorn in my hat then grass is green (understood as material implication) and 2) Grass is green, to 3) There is probably a unicorn in my hat, is obviously illegitimate. The conditionals we employ in ordinary discourse, however, are almost never to be understood as conditionals of material implication and it remains an open question as to whether the inference from 1) and 2) to 3) of argument (I) is a legitimate inference. I shall argue that it is not by first asking you to compare (I) with the following arguments:

(II) 1) There are footprints on the beach.
2) If Jimmy Carter had walked the beach recently, there

2.623 and 2.625). It also appears to reflect the form of the examples Harman gives of reasoning to the best explanation (1965 and 1968).

[6] I use the double line to indicate that the relation between premises and conclusion is only intended to be one of confirmation.

would be such footprints.

3) Jimmy Carter walked the beach recently.

(III) 1) There are footprints on the beach.
2) If a cow wearing shoes had walked the beach recently, there would be such footprints.

3) A cow wearing shoes walked the beach recently.

Now obviously, while (I) would seem to us a perfectly natural and legitimate case of reasoning to the best explanation, (II) and (III) (in most circumstances) would strike us as odd to say the least. But (I), (II) and (III) all have precisely the same form *and* true premises. Why is it that we accept (I) as a good argument to the best explanation while we reject (II) and (III)?

Since as we just noted we cannot isolate the difference by looking at the form of the arguments and since each argument has true premises, I would suggest that we accept (I) because (I) is an enthymeme. We accept a crucial but unstated premise from which we can legitimately infer the conclusion. That premise is the obvious one:

2a) In the vast majority of known cases footprints are produced by men.

We reject (II) and (III) because we do not accept the relevant implicit premise. It is not true that in the vast majority of known cases footprints on the beach are produced by Jimmy Carter, nor is it true that in the vast majority of known cases such footprints are produced by a cow wearing shoes. But now if 2a) is an essential part of the evidence from which we are willing to infer that a man walked the beach recently, it seems clear that the so-called argument to the best explanation is really just an inductive argument whose form would be put more perspicuously this way:

1) In all (or most) cases in which we have observed footprints on the beach there were men present just prior to the existence of such footprints.
2) Here is another case of footprints on the beach.

3) A man was present just prior to the existence of these footprints.

The argument is a standard inductive argument having the following form:

1) All or most of the A's we have observed were immediately preceded by B's.

2) This is an *A*.

3) It was preceded by a *B*.

The first premise asserts a discovered correlation between the occurrence of two states of affairs, the second premise asserts that one is present and the conclusion asserts that the other is present. It is true that in typical examples of inductive reasoning one moves from the observed correlation and the discovery of a fresh instance to the *later* occurrence of some state of affairs but that is clearly only because induction is so closely associated with the problem of knowing the future.

I have taken a very simple case of what seems to be a paradigm of reasoning to the best explanation and argued that though we might put the argument in ordinary discourse in a way that makes it appear superficially different from inductive reasoning, we only accept the argument as good because we implicitly accept a premise which plays a crucial role in our willingness to infer the conclusion. I have argued that this implicit premise is crucial by indicating that there are many arguments having true premises and exactly the same form as our intuitively acceptable case of reasoning to the best explanation but which are clearly unacceptable arguments. I have finally argued that when we make explicit the implicit premise of the argument to the best explanation, we have all of the premises we need to present a perfectly straightforward inductive argument.

In the course of presenting the above argument I claimed that we would obviously not accept (II) as a good argument. Let me qualify that by admitting that there are *circumstances* in which we might unreflectively present and accept (II) as a good argument to the best explanation. If, for example, I know that Carter and I are the only two people near a certain beach, I know that I haven't walked the beach recently, and I see footprints on the beach, I might reason to the conclusion that Carter walked the beach recently. Again, to the extent to which we might accept (II) as a good argument in such circumstances, I think it must be construed as an enthymeme, but the way in which we would fill out the implicit premise would differ somewhat from the previous case. As noted it is not true that in most cases footprints on the beach are immediately preceded by the presence of Jimmy Carter, so (II) cannot be reduced to a straightforward inductive argument. But we have the inductively supported conclusion that a man walked the beach recently (the conclusion of the inductive argument disguised as (I)) and from our independently supported conclusion that Jimmy Carter was the only man who could have walked the beach recently we can *deduce* that Jimmy Carter walked

the beach recently. Thus the reasoning underlying our acceptance of (II) in the hypothetical situation sketched above might be put more perspicuously as follows:

1) In all (or most) cases in which we have observed footprints on the beach there were men present just prior to the existence of such footprints.
2) Here is another case of footprints on the beach.

---

3) A man walked the beach recently.
1) A man walked the beach recently and Jimmy Carter is the only man who could have walked the beach recently.

---

2) Jimmy Carter walked the beach recently.

When we reason to a conclusion employing both inductive and deductive reasoning in the above manner, I shall refer to the process as indirect inductive reasoning.

I have argued then that if there are *circumstances* in which we might take (II) to be an acceptable argument to the best explanation we shall find that we only accept (II) as a good argument because we implicitly accept premises which again play a crucial role in our willingness to infer the conclusion and I have further argued that when we make explicit the implicit premises the reasoning can be seen to be indirect inductive reasoning. *Inductively* we reason to a certain conclusion and from other things we know conjoined with that inductive conclusion, we *deduce* the conclusion in question.

II

At this point one might complain that the conception of reasoning to the best explanation which I have been considering is too simplistic. Granted, there are an indefinite number of arguments having the form of (A), each of which has true premises, but part of what reasoning to the best explanation involves is careful employment of criteria for *choosing* between the (always indefinitely large number of) alternative possible explanations. One might argue that the following more adequately represents the nature of reasoning to the best explanation:[7]

(B) 1) *Q*
2) Of the set of available competing and incompatible hypotheses *P1, P2, . . ., Pn* capable of explaining *Q*, *P1* is the

[7]I am indebted to a referee of *Philosophy of Science* for suggesting this alternative conception of reasoning to the best explanation in comments on an earlier version of this paper.

best explanation of Q according to criteria *C1, C2, . . ., Ck*.

---

3) *P1*

To complete the above account of reasoning to the best explanation, one must, of course, indicate what these criteria, *C1, C2, . . ., Ck*, are. In a recent paper, Paul Thagard suggests that "actual cases of scientific reasoning exhibit a set of criteria for evaluating explanatory theories" (1978, p. 76). These criteria he identifies as consilience, simplicity and analogy.

According to Thagard, "one theory is *more* consilient than another if it explains more classes of facts than the other does" (1978, p. 79). (The use of "explains" here and in Thagard's other remarks is potentially misleading. Of a number of competing, i.e. incompatible theories there is a sense in which there is always only one that explains at all, at least if by "explains" we mean "correctly or adequately explains." When Thagard compares the way in which competing theories explain facts he means to be comparing the way in which these theories *would* explain facts if they were true.) Thagard emphasizes that it is *kinds* of facts explained rather than sheer number of facts explained that is the relevant test of the consilience of a theory and is quite open about the problem of indicating in any sort of precise way how we are to go about *counting* the number of kinds of facts explained by a theory.

Whether one explanatory theory is *simpler* than another, according to Thagard, is a function of the size and nature of the set of auxiliary hypotheses the theory needs to explain a range of facts, where by auxiliary hypotheses he seems to have in mind (roughly) *ad hoc* adjustments or additions to a theory for the specific purpose of insuring that the theory fits some particular range of facts.

Thagard introduces the concept of analogy with which he is concerned with an example:

> Suppose A and B are similar in respect to P, Q, and R, and suppose we know that A's having S *explains* why it has P, Q, and R. Then we may conclude that B *has* S is a promising explanation of why B has P, Q, and R (1978, p. 90).

One explanatory theory *T1*, then, will be better supported by analogy than another theory *T2* if we find that *T1* is analogous (in the above sense) to established explanations while *T2* is not.

In choosing between alternative explanations we should, according to Thagard, try to *balance* our criteria. We want, for example,

consilience, but not at the expense of simplicity; simplicity, but not at the expense of consilience.

Thagard's criteria for choosing between competing explanations are just examples of many that have been proposed. There have been, for example, countless variations on appeals to simplicity since Ockham offered his razor as a way of trimming ontological commitment.[8] Though I shall use Thagard's criteria as convenient examples, the following comments are designed to apply to any similar attempts to flesh out reasoning to the best explanation as an *alternative* to inductive reasoning.

The first thing we must do in assessing any such criteria for choosing the *best* of competing explanations is to make sure we are quite clear about the sense of "best" used. A theory which is more consilient and simple than alternatives (in Thagard's sense and in a number of other senses) is certainly more desirable than its competitors in the sense that it would be *nice* if it turned out to be true. In general, I assume we are interested in explaining as much as we can and a theory which explains a great deal, both in terms of number and kinds of facts, while avoiding unwieldy *ad hoc* additions, would be a happy theory to have. But this not being the best of all possible worlds (some theologians aside) what would be nice is not always so.[9] The relevant epistemological question is whether the more consilient, simpler, more analogous theory is, *ceteris paribus*, more likely to be true.

Moreover, in the present context our question becomes even narrower. Since we are concerned with the claim that reasoning to the best explanation constitutes an *independent* source of knowledge, we must determine whether we may know *a priori* that the more consilient, simpler, more analogous theories are, *ceteris paribus* more likely to be true. After all, if it were a mere contingent fact about the world, *revealed to us through experience,* then it seems clear that employment of Thagard's criteria would simply reflect a particular application of inductive reasoning. If, for example, we are justified in trusting more consilient theories, *ceteris paribus,* only because for the most part more consilient theories have turned out to be more successful, then inferring that the correct explanation is, *ceteris*

[8]For one very recent and somewhat novel discussion of simplicity, see (Fales, 1978). Though Fales does suggest that simplicity, in his sense, is a desirable feature for a theory to have, he does not explicitly claim that the simplicity of a theory makes it more likely to be true.

[9]Grover Maxwell emphasizes the same point in discussing simplicity as a criterion of truth (1975, pp. 159–160). I ought to feel somewhat embarrassed about enlisting Prof. Maxwell's aid here in attacking reasoning to the best explanation since (Maxwell 1975) is equally concerned with attacking the legitimacy of *inductive* reasoning.

paribus, the most consilient explanation would amount to a straightforward instance of reasoning by enumerative induction.

To defend the central thesis of this paper, then, I need not deny that Thagard's criteria are relatively successful in sorting correct from incorrect explanatory theories (I shall take no stand on this question), I need only make plausible the claim that *if* employment of such criteria is justified the justification is inductive, i.e. rests on our independently determining that for the most part correct theories satisfy these criteria.

Now it is difficult to establish conclusively that one sort of reasoning is parasitic upon another, but it seems to me that I can make my claim at least highly plausible by pointing out that *if* Thagard *were* correct there would almost certainly have to be an inductive justification for the employment of his criteria *available.* After all, if his criteria were successful then we may presume that correct explanatory theories *which deal only with observables* (i.e., which admit of direct inductive testing) display the properties of consilience, simplicity and analogy. If they did not, Thagard's thesis could be directly attacked through counterexamples. But if the correct explanatory theories which have been *inductively* established (i.e., established without having to appeal to criteria such as those proposed by Thagard) display the characteristics Thagard mentions, then we have an inductive justification for relying on such properties as indicators of truth when we are forced to make judgments with respect to more problematic sorts of theories. In short, either theories which have been established without appeal to criteria such as Thagard's[10] display the characteristics Thagard talks about or they do not. If they do not, his theory is surely discredited. If they do, we have available an *inductive* justification of the hypothesis that such criteria are reliable indicators of truth.[11]

Now to show that if Thagard's criteria are successful there will be *available* an inductive justification of their reliability is not to prove that reasoning employing such criteria is parasitic upon inductive reasoning. (After all, there is always available an inductive justification of the reliability of deductive inferences, but we wouldn't argue that

[10]The argument presupposes that there are at least *some* theories which can be established without having to employ the sort of criteria Thagard talks about. This one may wish to deny, though its denial goes far beyond anything Thagard defends in his paper. In fact, I think there is *some* plausibility to the claim that appeals to simplicity must *always* be involved in establishing a theory no matter how low-level the theory is. This does not affect my thesis, however, for, as I shall argue, there is, in many cases, a deductive justification for preferring the simpler of two theories.

[11]This argument is obviously closely related to the so-called pragmatic vindication of induction proposed by Reichenbach (1938) and Salmon (1970).

deduction is parasitic upon induction.) Nevertheless, by showing that we have no *need* of positing employment of such criteria as an *independent* kind of reasoning we have surely cast even more doubt on what must be an initially implausible view, namely that we can have *a priori* knowledge that theories exemplifying some set of properties, *C*, are more likely to be true than theories which do not.

I would like to add to the above remarks a few words concerning simplicity as a criterion for choosing between alternative explanations. Because so many philosophers seem to endorse appeals to simplicity it would seem unlikely that there is nothing to such appeals. As I have argued above it is certainly possible that there is an *inductive* justification for regarding the simplicity of a theory as an indication of its truth. In addition, it should be noted, there is, in certain circumstances, a *deductive* justification for preferring the simpler of two theories. Suppose I am considering two incompatible theories *T1* and *T2* which, relative to my evidence, are equally likely to be true. Suppose, further, that after acquiring some additional evidence (let us call my new *total* body of evidence *E*) I find it necessary to add an hypothesis *H1* to *T2*. Now provided that the probability of *T1* and *T2* relative to *E* remains the same and assuming that the probability of *H1* relative to *E* is less than 1, we can *deduce* from the probability calculus that, relative to *E*, *T1* is more likely to be true than the more complex theory (*T2* and *H1*). Intuitively, *T2 by itself* ran the same risk of error as *T1*, so with the addition of another hypothesis which might be false it runs a *greater* risk of error than *T1*.

As I had occasion to note in Part **I**, in arguing that reasoning to the best explanation collapses into inductive reasoning, I am certainly not overlooking the role that deduction plays in helping to sort through inductively supported theories, and I suspect that a purely formal justification for preferring a simple theory to a more complex theory (*ceteris paribus*) is often available.

III

I expect that one might object to the arguments of Part **I** by claiming that I have seriously prejudiced the issue by focussing on such simplistic cases of reasoning to the best explanation. Perhaps we can reduce reasoning to the best explanation to inductive reasoning in such straightforward cases (where after all there is no difficulty observing the relevant correlations) but it is something else to argue that the scientific inference to the existence of certain *unobservable* entities based on complex observation of sophisticated instruments can so easily be reduced to inductive reasoning.

To this I have two replies. First as far as I am concerned the reductionist-instrumentalist-realist[12] debate is still a live issue in the philosophy of science. We should not choose our epistemic principles with the realist's ontological commitments as our guiding light. We should test our intuitions about what does or does not constitute legitimate reasoning by looking at just the sort of cases I have considered in Part I. If in all of the commonplace arguments to the best explanation we employ and accept, we see that such arguments must contain, at least implicitly, a premise which allows us to reduce the argument to an inductive argument and if the scientific realist cannot come up with the required premise, then perhaps we should look for another way of understanding his statements about these troublesome entities, a way of understanding such statements compatible with our being able to justify our belief in them.

Secondly, it is not at all clear that employing only inductive reasoning, the scientific realist cannot justify his beliefs in many of the propositions asserting the existence of the sorts of things to which he wants to give full-bodied existence. Philosophers too often forget that inductive reasoning itself can become extraordinarily complex. Reaching a conclusion inductively based on one's *entire* body of evidence usually involves countless inductive arguments the premises of which confirm countless hypotheses to varying degrees of probability, some of which hypotheses are compatible, some incompatible, with one another. What one is justified in believing with such an evidential framework depends on the result of weighing a great many probabilities. Though I don't really know, I suspect that the scientist has perfectly good *inductive* evidence justifying his belief in the existence of many of the theoretical entities to which he is committed (where the existence of such things is understood precisely as the realist would have us understand it). *Though theoretical entities are themselves unobservable, their defining properties may be such that we have observation of other things having those properties.* This opens the door to the kind of indirect inductive reasoning sketched earlier.

In any event the reductionist-instrumentalist-realist dispute should be judged case by case. The instrumentalist and reductionist are right in insisting that the realist be prepared to come up with inductive

[12]The terms "reductionism" and "instrumentalism" are used in a number of different ways by philosophers. As I use the term, the "reductionist" claims that theoretical statements are analytically equivalent to statements about the macroworld (usually conditionals describing how certain macroentities would behave under certain observable conditions). The reductionist, then, like the realist, takes theoretical statements to make assertions and thus to be straightforwardly true or false. The instrumentalist on the other hand thinks of talk about theoretical entities as a kind of convenient fiction. Such talk may be judged useful or not, but may not be judged straightforwardly true or false.

justification or choose between skepticism and either instrumentalism or reductionism. In some cases the realist will be able to find inductive support; in some cases he will not and will find the instrumentalist's "convenient fictions" the most attractive alternative; in others, he will not and will find the reductionist's analysis the most attractive alternative; in still other cases he may be forced to simply abandon his belief.

The traditional problem of perception, incidentally, is a different matter. If reasoning to the best explanation really does collapse into inductive reasoning (or even if inductive reasoning underlies reasoning to the best explanation), then one still has to tell Hume how to get beyond his sensation to a world logically but not causally independent of them. The sort of complex inductive reasoning discussed above does not seem available to us at this early stage of our attempt to reconstruct the foundations of empirical knowledge for not having gone beyond sensation yet, it is not clear how anything we can correlate would be in any way relevant to establishing the existence of physical objects. Here I suspect that as long as one works within the framework of radical empiricism, one is faced with the extremely difficult task of choosing between Humean skepticism and the complexities of a phenomenalistic analysis.

REFERENCES

Ennis, Robert H. (1968), "Enumerative Induction and Best Explanation," *The Journal of Philosophy 65*: 523–529.

Fales, Evan (1978), "Theoretical Simplicity and Defeasibility," *Philosophy of Science 45*: 273–288.

Harman, Gilbert H. (1965), "The Inference to the Best Explanation," *Philosophical Review 74*: 88–95.

Harman, Gilbert H. (1968), "Enumerative Induction as Inference to the Best Explanation," *The Journal of Philosophy 65*: 529–533.

Hume, David (1888), *A Treatise of Human Nature,* L. A. Selby-Bigge (ed.), Oxford: Clarendon Press.

Locke, John (1959), *An Essay Concerning Human Understanding,* Alexander Fraser (ed.), New York: Dover.

Mackie, J. L. (1969), "What's Really Wrong with Phenomenalism," *The Proceedings of the British Academy 55*: 113–127.

Maxwell, Grover (1975), "Induction and Empiricism: A Bayesian-Frequentist Alternative." In *Minnesota Studies in the Philosophy of Science,* Vol. VI, Maxwell and Anderson (eds.), Minneapolis: University of Minnesota Press.

Peirce, C. S. (1931), *Collected Papers,* C. Hartshorne and P. Weis (eds.), Cambridge: Harvard University Press.

Reichenbach, H. (1938), *Experience and Prediction.* Chicago: The University of Chicago Press.

Russell, B. (1959), *The Problems of Philosophy.* London: Oxford University Press.

Salmon, Wesley C. (1970), "Inductive Inference." In *Readings in the Philosophy of Science,* Baruch A. Brody ed., Englewood Cliffs: Prentice-Hall: 597–612.

Thagard, Paul (1978), "The Best Explanation: Criteria for Theory Choice," *The Journal of Philosophy 75*: 76–92.