EPISTEMIC MERIT, INTRINSIC AND INSTRUMENTAL*

Roderick Firth
Harvard University

Philosophers today are deeply divided by issues of a kind that can be described as "metaepistemological" -- issues concerning the concepts and methodology appropriate for a philosophical investigation of human knowledge. In thinking about some of these controversial issues it seems to me helpful to draw a distinction between two kinds of epistemic merit – intrinsic and instrumental. There are many things that can possess epistemic merit: educational practices, methods of scientific inquiry, moral and religious doctrines, constitutional provisions for freedom of speech and inquiry, and so on indefinitely. But I shall be talking today about the epistemic merit of propositional attitudes; and for simplicity I shall concentrate attention on just one of these, the attitude we have toward a proposition in believing it. What I shall have to say about the epistemic merit of believing can easily be extended to other attitudes like disbelieving, suspending judgment, assuming for the sake of the argument, accepting subject to further investigation, and so on.

Now let us suppose that John Doe is suffering from an illness that is usually fatal, but believes with deep conviction that he will recover. The fact that John has this optimistic belief might actually contribute to his recovery. Or at least it might make him more cheerful during his dying days, which in turn might ease the pain of others who are close to him. In either case it

---

*Presidential Address delivered before the Seventy-Seventh Annual Eastern Meeting of the American Philosophical Association in Boston, Massachusetts, December 29, 1980.

would be a virtue or *merit* of John's belief that it has the good consequences that it does for himself or others. The merit of the belief would be instrumental, varying in degree with the value of the good consequences. Merit of this kind might also be classified as ethical. For suppose that John had, as we say, *made* himself believe that he would recover – either by commonplace devices or by sophisticated psychological exercises – so that he is morally responsible for his belief. In that case we might conclude, after weighing all the relevant considerations, that the instrumental merit of the belief is so great that John is ethically justified in believing that he will recover. He is ethically justified in believing that he will recover just as he is sometimes ethically justified in doing other things that reduce his own suffering and the suffering of others.

So far, however, the example is not one of *epistemic* merit, the merit that a belief may have because of some suitably close relationship to *knowledge*. But it is evident that a belief may be causally related to knowledge just as John's belief is causally related to the reduction of human suffering. Many beliefs, whether true or false, contribute causally to the growth of knowledge. Such beliefs, we may say, have instrumental epistemic merit. They might be general metaphysical assumptions of the kind that have been called "heuristic" -- beliefs, for example, to the effect that the universe is orderly in specified respects. But they might be beliefs that are narrow and personal. Thus if John were engaged in important research, his belief that he will recover from his illness might have instrumental epistemic merit. It might encourage him to continue working and it might thus contribute to the growth of knowledge.

Beliefs that contribute causally to the growth of knowledge have instrumental epistemic merit in the strictest sense of the term 'epistemic'. Their epistemic merit lies in the fact that they promote the growth of nothing less than *knowledge*, however difficult it may be to determine just what knowledge is. In the present context, however, it is convenient to extend this narrow use of the term 'epistemic' in two ways. Let us stipulate, first, since true belief is a necessary condition of knowledge,

that any belief, whether itself true or false, has instrumental epistemic merit to the extent that it contributes causally to the growth of our stock of *true beliefs*. And let us stipulate, second, that any belief has instrumental epistemic merit to the extent that it contributes causally to our stock of *warranted* beliefs, beliefs, that is to say, that are based rationally on the relevant evidence or which derive commensurable rationality from some other source. The degree of instrumental merit earned by a belief in these two ways would be decreased proportionately by the extent to which it produces false and unwarranted beliefs; but this is a complication that does not affect the issues I shall be discussing. For simplicity, then, let us say that a belief can earn instrumental epistemic merit in any or all of three causal roles: (i) as a means by which we acquire true beliefs, (ii) as a means by which we acquire warranted beliefs, and (iii) as a means by which we acquire knowledge. If, as suggested in the *Theaetetus*, knowledge entails warranted belief as well as true belief, then a belief that promotes knowledge will play all three roles at once. But those of us who reject that ancient doctrine are free to hold that, other things being equal, a belief acquires more epistemic merit by promoting what we might call "warranted knowledge" than by promoting knowledge that is unwarranted.

How is instrumental epistemic merit related to ethical merit?

This question brings up controversial issues, but we can probably agree on at least this much: a belief with instrumental epistemic merit, i.e., a belief that promotes true beliefs, warranted beliefs, or knowledge, will share derivatively in whatever ethical merit is possessed by the beliefs and knowledge it promotes. If a particular body of medical knowledge will decrease human suffering, and this fact makes such knowledge ethically desirable, a belief that promotes that medical knowledge will thereby derive some degree of instrumental ethical merit. But this is not the end of the matter, for many moral philosophers have held that truth, rationality, and knowledge are ends in themselves – which means, in effect, that a belief with instrumental epistemic merit is *ipso facto* a belief with instrumental ethical merit. Fortunately we can set aside these ethical issues for present

purposes by using the term 'epistemic merit' in a way that is ethically neutral. Thus we may even say, without logical inconsistency, that some beliefs have *too much* epistemic merit. We may argue that they promote human rationality or human knowledge to an extent that is ethically undesirable.

It is now easy to see how I propose to distinguish the *intrinsic* epistemic merit of a belief from the three kinds of instrumental epistemic merit just enumerated. A belief has intrinsic epistemic merit if one or more of these three conditions obtains: (i) the belief constitutes knowledge, (ii) the belief is true, (iii) the belief is warranted. As before, nothing is implied about either the epistemic or the ethical weights to be attached to the merit earned under each of these three conditions. We are free to consider whether in some or all situations warranted knowledge has more intrinsic *epistemic* merit than the same knowledge without warrant, and whether beliefs acquire any intrinsic *ethical* merit from being warranted or from being instances of knowledge. For present purposes it is only important to recognize that intrinsic and instrumental epistemic merit may vary quite independently of one another. There are circumstances in which false beliefs may lead causally to true beliefs, warranted beliefs may lead to false beliefs, and so on for all the possible ways in which beliefs with and without intrinsic epistemic merit might produce other beliefs with and without epistemic merit. Consider, to take just one dramatic example, the belief that human beings are created and sustained by God for the primary purpose of expanding human knowledge. This belief might have no intrinsic epistemic merit at all. It might not be an instance of knowledge. It might not be true. And it might not be warranted. Yet it might have a very high degree of instrumental epistemic merit. If the belief were a common one, it might alter human motivation in a way that would produce undreamed of advances in human knowledge.

Now this distinction between two kinds of epistemic merit, intrinsic and instrumental, may seem to be a very obvious one, not one that philosophers are likely to overlook. But the distinction makes a number of basic epistemic terms ambiguous

in a way that can easily lead to confusion. Epistemologists often say, for example, that they are attempting to identify the conditions under which beliefs are *justified*. To avoid one ambiguity, already discussed, they may explain that they are talking about *epistemic* justification, not ethical justification. But another ambiguity remains. For what does it mean to say that a belief is epistemically justified? Does it mean that the belief has one of the three specified kinds of intrinsic epistemic merit -- the kind, namely, that is characteristic of beliefs that are supported by the relevant evidence? Or does it mean that the belief has instrumental merit in virtue of its epistemic consequences? In either case it seems appropriate to say that the belief is *epistemically* justified -- or, alternatively, that it is epistemically reasonable, rational, or desirable. But the ambiguity may obscure the fact that there are two quite distinct epistemological tasks that we could have in mind when we say that we are trying to identify the conditions under which beliefs are epistemically justified.

Each of these tasks is represented by a long philosophical tradition. On the one hand we may be trying to formulate rules or instructions for distinguishing what I have called "warranted" beliefs from unwarranted beliefs. Some of these rules will be rules for identifying those kinds of things that are capable of serving as evidence for our beliefs - perhaps sense experiences, perhaps some select subclass of our current beliefs, and so on. Some of them will be rules for distinguishing beliefs that are supported by a given body of evidence from beliefs that are not supported by that evidence. So we may conveniently say that our task is to work out a theory of evidence. A complete theory of evidence would propose necessary and sufficient conditions for warranted belief. On the other hand we may be attempting another epistemological task that has an equally honorable tradition. We may be attempting to formulate rules for identifying beliefs with *instrumental* epistemic merit, beliefs that promote the advancement of knowledge or some other epistemic end. Rules for choosing between competing scientific hypotheses are examples of such rules, provided that issues concerning warrant are excluded by assuming that the competing

hypotheses are equally warranted on the basis of the available evidence. To the extent that we thus seek to evaluate the instrumental epistemic merit of beliefs – or, indeed, of anything else – we are engaged in calculations of epistemic utility; and without denying that there are other epistemological tasks, we may call ourselves "epistemological utilitarians." As epistemological utilitarians we must study the cognitive mechanisms and procedures that have led in the past to the acquisition of true beliefs or knowledge, and we must spot defects in them that have led to false beliefs. We can then draw inductive conclusions in a way characteristic of the natural sciences. The epistemological utilitarian is engaged, therefore, in the kind of investigation that Quine has aptly called "naturalized epistemology."

The task of the epistemological utilitarian is a complex one, embracing large areas of psychology, sociology, biology, and other sciences. But whatever the difficulties, it seems not to raise any *methodological* problems beyond those already raised by scientific inquiry in general. The task of the epistemological utilitarian is simply the task of identifying certain causal relations, and is thus one of the tasks embraced by the broad term 'natural science'. There are legitimate philosophical motives, therefore, for trying to reduce questions at issue in theory of evidence to questions of epistemic utility. There is a close analogy here to the attempts of ethical utilitarians to reduce questions of right and wrong, justice and injustice, praiseworthiness and blameworthiness, to questions of ethical utility. And in both cases the reduction may be either criteriological or conceptual. Epistemological utilitarians might maintain that an adequate *criterion* for the warrant of a belief can be formulated without circularity in terms of instrumental epistemic merit. Or they might maintain more than this, namely, that the criterion is adequate in virtue of a *conceptual* relationship between warrant and epistemic utility.

Such a conceptual reduction might be similar, for example, to G. E. Moore's utilitarian analysis of the concept of right. Although he later repudiated the doctrine, Moore maintained in *Principia Ethica* that to say that an act is right is simply to

say that there is no other act open to the agent that would have better consequences. This reduces rightness to instrumental ethical merit, and has the effect, so to speak, of ruling out deontological positions by definition. A similar conceptual reduction of epistemic warrant might be proposed by epistemological utilitarians. To say that a belief is warranted, they might maintain, is to say something about the superior epistemic merit of the causal consequences of that belief as contrasted with the epistemic merit of the causal consequences of alternative beliefs of some specified kind. This reduces warrant to instrumental epistemic merit, and has the effect of converting all questions concerning theory of evidence into questions about epistemic utility. And it should not be overlooked that this result can be achieved indirectly by definitions of warranted belief that involve conative or ethical terms in a vacuous way. Suppose, for example, that the warranted belief in a specified situation is defined as the belief that would best satisfy the desires of someone who desired only to maximize true beliefs. Or consider an ethical variation: the warranted belief is the one that would best fulfil a *duty* to maximize true beliefs. If maximizing true beliefs means maximizing true beliefs throughout future time, such conative and ethical definitions, whether so intended or not, reduce warrant to epistemic utility.

Now the traditional kind of ethical utilitarianism advocated by Moore, whether construed as a conceptual theory or simply as a criteriological theory, is the kind that is today often called act-utilitarianism. The question whether a particular act is right or wrong is reduced to a question about the instrumental merit of that act itself. And whatever we may think of this position, it seems clear that no such simple reduction is plausible in the case of epistemic warrant. As my examples have shown, a belief that is *not* supported by the relevant evidence might have great instrumental epistemic merit, and a belief that *is* supported by the evidence might have none at all. There just is no guarantee that the warrant of a particular belief is proportionate, or even roughly proportionate, to its efficacy in producing knowledge or true beliefs in the long run. Thus the epistemological analogue

of ethical act-utilitarianism, a position that we might call epistemological belief-utilitarianism, is not a position that anyone is likely to advocate.

What shall we say, however, about the plausibility of epistemological positions that are analogous to ethical *rule*-utilitarianism? Such positions are sometimes called reliability theories of justification, but might well be described as forms of *epistemological* rule-utilitarianism.* All the various possible forms of ethical rule-utilitarianism have some claim to being called utilitarianism, for they all reduce questions about the rightness and wrongness of an act to questions about ethical utility. But they do so by asking us, so to speak, to shift our attention from the long-run consequences of a particular act to the long-run consequences of choosing our acts in accord with certain rules. Such theories can specify the relevant population of choosers and the conditions of choice in very different ways, but they all allow an act with ethically inferior consequences to be right if it is chosen in accord with a set of rules that has ethical utility in the long run. What is demanded of the set of rules, roughly speaking, is just that it be statistically *reliable* to a degree specified by the theory. Acting in accord with the rules, under appropriate conditions, must *tend* to maximize consequences of a kind that the theory identifies as good.

Epistemological rule-utilitarianism has a similar claim to being called utilitarianism. In deciding whether a particular belief is warranted, it asks us to consider not the instrumental merit of that particular belief, but the long-run consequences of arriving at our beliefs, under specified conditions, in ways that accord with certain rules. Thus a belief with no instrumental epistemic merit may nevertheless be warranted because it is acquired in accord with rules that have epistemic utility in the

---

* A reliability theory of justification, it should be noted, is to be distinguished from a reliability theory of knowledge. It may be necessary for an adequate account of knowledge to make some kind of appeal, direct or indirect, to reliability, whether or not this is necessary in the case of justification.

long run. The rules might stipulate ways of collecting evidence, ways of assigning weights to various kinds of evidence, ways of drawing inferences, and so on. What is demanded of the set of rules is, roughly speaking, that it be instrumentally reliable -- statistically reliable as a guide to something with intrinsic epistemic merit. It is this kind of reliability that makes the rules "valid" and thus able, so to speak, to confer warrant on a belief. If we are to avoid circularity, the end to be promoted cannot be defined by reference to states of warranted belief. But it might be defined by reference to states of *true* belief. And, if knowledge does not entail warranted belief, the end to be promoted might be states of knowledge. To simplify discussion let us assume that the end is to produce true beliefs.*

This still leaves us with an ambiguity, however. There are two very different ways in which a set of rules might be said to lead reliably to true beliefs. Following the rules might tend to produce true beliefs *immediately*, so that we usually acquire a true belief each time we do so. Or the rules might tend to produce true beliefs in the long run, perhaps in part by producing immediately certain false beliefs that have long range utility. I have suggested a theological belief, for example, that might alter human motivation in a way that would have this effect. We can easily think of others. We can also easily think of rules of inference which, if generally followed, would produce beliefs of this epistemically useful kind. When their range is carefully restricted, the addition of such rules would increase the epistemic utility of any set of rules that tends to produce true beliefs immediately. Following the expanded set of rules would increase the proportion of true beliefs in the long run. But of course these additional rules would not be rules by which we ordinarily test the warrant of a belief. Whatever their ultimate utility,

---

*For a sophisticated version of what I call epistemological rule-utilitarianism, see Alvin I. Goldman, "What is Justified Belief," in George Pappas, ed., *Justification and Knowledge*, Boston, Dordrecht, London, 1979. I am much indebted to this essay, which gave rise to some of the following reflections.

therefore, I shall assume that they are not to be included among the rules that an epistemological rule-utilitarian can plausibly recognize as warrant-conferring. Reliable warrant-conferring rules are to be rules that tend to produce true beliefs immediately -- which is not to deny, of course, that in this way they may also increase the proportion of true beliefs in the long run.

This is not the place to attempt an evaluation of *ethical* rule-utilitarianism. But one objection to this ethical doctrine is that it makes the rightness or wrongness of an act dependent in an implausible way on remote past and future events not causally related to the act itself. And this draws attention to a more serious objection to epistemological rule-utilitarianism if construed as an analogous theory, i.e., as an attempt to provide us with *defensible* criteria for identifying warranted beliefs. For according to epistemological rule-utilitarianism, the degree of warrant attached to the belief of a particular person at a particular time, is positively correlated with the degree of statistical reliability overall of some set of rules for arriving at true beliefs. And however the rules are conceived, this implies that *any* series of events that affects the statistical reliability of the rules thereby affects the degree to which the particular belief is warranted. The emphasis here is on the word 'any'. For if we are talking about what I have elsewhere called "doxastic warrant,"* the warrant of a state of believing as contrasted with the warrant of the proposition believed, it is undeniable that *some past* events are relevant to the warrant of a belief. To decide whether someone is doxastically warranted in believing a particular proposition, we must know what processes -- what "belief-generating processes," as we may call them -- produced or sustained the belief. For it is on these processes that we are ultimately passing judgment. But this fact is -- or at least should be -- neutral among epistemological theories.

Let us suppose, to make the issue more concrete, that our John Doe is *not* doxastically warranted in believing that he will

----

*"Are Epistemic Concepts Reducible to Ethical Concepts," in *Values and Morals,* edited by A. I. Goldman and J. Kim, 1978, D. Reidel, Holland, pp. 217-19.

recover from his illness. Let us suppose that he commits a form of the Gambler's Fallacy in reaching and maintaining his optimistic belief. Doctors often make mistakes, he thinks, but throughout my lifetime their diagnoses and prognoses of my illnesses have invariably been correct. So it is now time, he reasons, for the doctors to make a mistake; and fortunately for me they now predict that my present illness is fatal. In neutral terminology we are supposing that John's belief is thus the result of processes that are not epistemically "valid." But according to epistemological rule-utilitarianism, these belief-generating processes *would* be valid, however we now evaluate them, if only they were in accord with statistically reliable rules for arriving at true beliefs. And statistical reliability, of course, is determined by the actual course of events overall, past, present, and future. (This is true whether the reliability of the rules is to be determined by the percentage of actual past, present, and future cases in which following the rules leads to true beliefs, or by the percentage of cases, past, present, and future, in which following the rules *would* lead, or *would* have led to true beliefs, whether or not the rules were actually followed. For what would happen if someone were to follow the rules at a particular time *t* depends on other causal conditions that actually obtain before and after *t*.) If we accept this reliability theory of warrant, therefore, we must accept the conclusion – among others almost as startling -- that the truth of our judgment that John has reasoned fallaciously depends on the course of future events. For according to this theory John's belief would be warranted, however he arrives at it, if it should happen that there is a sufficiently long period in the future, whether or not we live to know about it, when belief-generating processes like John's do (or would) yield true beliefs with a suitably high frequency. The fact that epistemological rule-utilitarianism commits us to such a conclusion seems to me to make that theory rationally indefensible in a way that needs to be carefully defined.

When I say that epistemological rule-utilitarianism is rationally indefensible, I do not intend to suggest that the theory is false. We are all convinced, I am sure, that beliefs acquired in accord

with valid epistemic rules, i.e., rules that confer warrant, tend to be true, and that beliefs acquired by fallacious reasoning, as was John's, do not tend to be true. We are convinced that this correlation between warrant and epistemic utility existed in the past and will continue to exist in the future. This conviction, indeed, or at least behavior compatible with this conviction, is partly definitive of the character trait that we call rationality in epistemological contexts. We may not want to say that a rational person, in this sense of the word, puts truth categorically above all other goals. But a rational person is always someone whose belief-generating processes are compatible with the following *hypothetical* imperative: If you want to believe a particular proposition if and only if it is true, believe it if and only if it is warranted. Thus no rational person can believe simultaneously (1) that John Doe's belief fails to be warranted because based on fallacious reasoning, and (2) that John's method of inference is overall a statistically reliable guide to true beliefs. This is not to say, however, that a rational person can have a *reason* to believe that there is a correlation between warrant and statistical reliability. If we could think of any such reason, then that would be a reason for us to accept epistemological rule utilitarianism. I should be wrong in asserting that the theory is rationally indefensible.

It is no doubt paradoxical to say that as rational people we all believe that there is a positive correlation between warrant and statistical reliability, and yet to deny that we have any reason to believe this. For when we say that rational people believe such and such, we usually imply that there are good and obvious reasons to believe such and such. In this case, however, it is just the believing that matters, not reasons to believe. Consider, by way of analogy, what we mean when we say that someone is a *trusting* person. We mean, at least in part, that the person believes that in general other people are truthful, honest, faithful, and so on - or at least that the person behaves in a way compatible with believing these things. To be trusting we need not have a *reason* to believe these things. It is enough that we believe them. And so, to be rational, we need not have a reason to

believe that warrant is a statistically reliable index of truth. We might even say, indeed, exploiting the analogy, that the rational person is in this respect a trusting person. The rational person *trusts* that there is a positive correlation between warrant and truth, and acts accordingly.

But is it true, as I have said, that there is actually no reason to think that this correlation exists, and hence no reason to accept epistemological rule-utilitarianism? Could it be shown, perhaps, that the correlation between warrant-conferring rules and true beliefs rests on conceptual grounds? Could we argue, for example, that to judge that John Doe's belief is unwarranted, is precisely to judge that his belief was not produced in accord with statistically reliable rules for generating true beliefs? But this would imply, counterintuitively, that in asserting that John's belief is unwarranted, or that his belief is based on fallacious reasoning, we are quite literally *predicting* the future. We are predicting that the course of future events will not be one such that, in conjunction with the course of past events, John's mode of inference is a statistically reliable guide to truth. What could persuade us to accept this implausible analysis of our meaning? The assertion that John's belief was unwarranted seems as little like a prediction as the assertion that John believed that he was ill.

It should be noted, moreover, that if this *were* a correct analysis of our concept of warrant, and judgments about warrant *were* judgments about statistical reliability, a new problem would arise. The advantage of this proposed analysis is that it allows us to establish the reliability of warrant-conferring rules without using an inductive argument. By 'warrant-conferring rules', it is said, we *mean* statistically reliable rules. But in that case is there any way to decide rationally whether or not a particular set of rules, e.g., those followed by John Doe, is or is not warrant-conferring? To decide this, which involves predicting the future, would require an inductive argument based on the past failure or success of those rules in generating true beliefs. But if we are rational, how can we identify true beliefs except by assuming, explicitly or implicitly, that some particular set of rules is warrant-

conferring? This is a special form of the central problem that faces us if we try to show that epistemological rule-utilitarianism can be defended by inductive argument. So let us consider the central problem as it arises in that context.

Since this epistemological theory is a form of utilitarianism, it is natural to suppose that it might be established, like causal theories in general, by inductive inference. Thus *ethical* rule-utilitarians try to defend their position by showing that those acts that we believe pretheoretically ("intuitively") to be right are the very ones that would be picked out by the general rule-utilitarian criterion. They try to show, in other words, by appealing to causal laws, that those acts are the very ones that *would* be chosen if we were to choose in accord with a set of rules that in fact maximized good consequences. An analogous defense of *epistemological* rule-utilitarianism would require an inductive causal argument linking truth to those beliefs that we consider pretheoretically ("intuitively") to be warranted. Such beliefs, so the argument would run, are the very beliefs that we should have if our beliefs were generated in accord with a set of rules that in fact maximized true beliefs. But such an argument presupposes that it is possible to identify true beliefs in a way that is independent of the way we identify warranted beliefs pretheoretically. If we make the crucial assumption that this is possible, perhaps we could argue that from past observation we have learned that true beliefs have been generated in accord with a certain set of rules, rules that might therefore be called truth-eliciting rules. Then we might argue that from past observation we have also learned that warranted beliefs are generated in accord with those very same rules. And from this we might conclude that rules tending to lead us to true beliefs constitute an adequate criterion of warrant. This would be analogous to the position of the ethical rule-utilitarian who holds that it can be shown by inductive generalization that rules of action that tend to lead to good consequences constitute an adequate criterion of ethical rightness.

This seems like a possible line of argument for epistemological rule-utilitarianism, however, only until we consider the crucial

assumption that true beliefs (and hence truth-eliciting rules) can be identified independently of warranted beliefs (and hence warrant-conferring rules.) Beliefs do not come labeled true and false so that we may discover inductively how true beliefs are correlated with warranted beliefs and warrant-conferring rules. To the extent that we are rational, each of us decides at any time *t* whether a belief is true, in precisely the same way that we would decide at *t* whether we ourselves are, or would be, warranted at *t* in having that belief. Whether the method we follow is valid or invalid, we are irrational if we have a *t* one way of assessing warrant and another way of assessing truth. If we are rational we must assume, either implicitly or explicitly, a correlation between warrant-conferring rules and true beliefs -- as, indeed, we all do – in order to identify true beliefs. Our reasoning is obviously circular if we then use beliefs so identified as data for inferring a correlation between warrant-conferring rules and true beliefs. The circularity guarantees that we shall find the correlation.

This circularity could be avoided if we abandoned the attempt to defend epistemological rule utilitarianism as a *general* theory about the conditions of warranted belief, and argued merely that *some* warrant-conferring rules can be identified by inductive inference. Some rules involving an appeal to *authority*, for example, can be so identified without circularity. We can argue from past experience to the conclusion that people of such and such a kind usually give true answers to questions about such and such subjects. We can do this, of course, only because we have an independent criterion of truth, an independent way of determining that the answers have been true in the past. But there seem to be certain rules or principles of inference – e.g., a rule for distinguishing warranted from unwarranted memory beliefs, and an inductive rule (violated by the Gambler's Fallacy) for generating general conclusions from particular cases -- that *must* be employed if we are to show inferentially that any rule whatsoever is a reliable guide to true beliefs. And such "basic" rules, as we may call them, are the ones that cannot be identified without circularity as reliable guides to true beliefs.

Even with respect to such basic rules the circularity could be avoided if each of us were to construe epistemological rule-utilitarianism as a theory about the warrant of the beliefs of *other* people, thus placing our own beliefs and questions about their warrant, beyond the range of the theory. In principle this would allow *me*, for example, to discover inductively which rules, including basic rules, are reliable guides to true beliefs (and thus warrant-conferring) when followed by other people. For I should not have to assume, in circular fashion, the universal validity of any of the rules I myself follow when distinguishing the true beliefs of other people from their false beliefs. But if we relativize epistemological rule-utilitarianism in this way, I, like the rest of us, am still left with an epistemological question that even the solipsist cannot avoid: Under what conditions are *my* beliefs warranted? And any argument for extending the criterion of reliability to determine the warrant of my own beliefs, would involve the circularity we have been considering.

This circularity, it should be noted, is not the circularity that so often characterizes attempts to justify inductive generalization. It is a circularity that infects arguments in support of epistemological rule-utilitarianism even if no question at all is raised about the validity of inductive generalization. If is a special kind of circularity, furthermore, that distinguishes epistemological rule-utilitarianism from other theories of evidence that propose necessary and sufficient conditions for warranted belief. It can be argued that in one respect all such theories of evidence are based on circular reasoning. For in order to present arguments in support of a particular theory we must start from what we take pretheoretically to be instances of warranted and unwarranted beliefs and draw theoretical conclusions from these instances. But this means that we must assume from the outset the validity of whatever modes of non-deductive inference we use to draw our theoretical conclusions -- modes of inference that would not be acceptable on some conceivable theories of evidence. How to respond to this charge, and its skeptical implications, is a problem shared by all such theories. But the special problem of circularity faced by the epistemologi-

cal rule-utilitarian is an *additional* problem. It arises when we ask how a rational person can identify true beliefs as a preliminary step to deciding by induction whether there is a correlation between true beliefs and beliefs generated in accord with warrant-conferring rules.

This suggests that one way to summarize what I have said about epistemological rule-utilitarianism is to maintain that it is not a genuine theory of evidence. If this seems too arbitrary, however, we might say that it is at best a *neutral* theory, one that any rational person must accept, and in that respect not rationally debatable. Thus it is not in competition with debatable theories of evidence such as coherence theories, traditional forms of foundationalism, and intermediate combinations. Debatable theories like these do not of course deny that belief-generating processes confer warrant only if they tend to produce true beliefs. Any theory that denied this would be unacceptable to a rational person. Theories of evidence that are rationally debatable propose answers, in effect, to the question: Under what conditions do belief-generating processes do *both* of these things. They propose answers directly to the question: "Under what conditions do belief-generating processes produce warranted beliefs?" And, as any rational person must agree, they *thereby* propose answers -- the same answers - to the question: "Under what conditions do belief-generating processes tend to produce true beliefs?"

I conclude, therefore, that there is no way of reducing the questions to be answered by a theory of evidence to questions of epistemic utility. But this is not to deny the independence and importance of the kind of inquiry undertaken by epistemological utilitarians who seek to formulate rules for the achievement of epistemic goals in the long run. Psychologists, sociologists, biologists, and other scientists might point out that there is nothing typically philosophical about an investigation of the causal factors that are conducive to the acquisition of true beliefs or the advancement of knowledge. Why not leave this investigation to us, they might ask. But they could scarcely deny that it is scholars trained in philosophy, usually identified as

philosophers of science, who are responsible for most of the work that has been done on the relative merits of various ways of choosing among competing scientific hypotheses. Within the broad field to be investigated by the epistemological utilitarian, this particular piece of territory seems to belong to philosophers by right of occupation. For this very reason, however, special care is needed if we are to avoid confusing questions about epistemic utility – questions about rules for achieving long-run epistemic goals – with traditional philosophical questions about the conditions of warranted belief.

We might well pledge ourselves, for example, never to ask general questions like "What characteristics can make one of two competing hypotheses epistemically more acceptable than the other?" or "What epistemic rules should guide our choice among competing hypotheses?" For if we are interested in the intrinsic epistemic merit of the hypotheses -- their warrant, their truth, and their eligibility for the status of knowledge – the answers to these questions call for a theory of evidence. But if we are interested in instrumental epistemic merit – the long-run epistemic advantages of preferring one hypothesis to the other – then many other considerations might be relevant: perhaps their fruitfulness in suggesting new experimental investigations, the extent to which accepting them would require revision of current scientific doctrine, and – on some interpretations of this controversial term – their simplicity.* It is an interesting question, indeed, why the epistemological utilitarian should pay any attention at all to the warrant and truth of the competing hypotheses. Has it ever been shown – indeed, *can* it be shown -- that warranted or true beliefs tend to have greater instrumental merit in the long run that unwarranted or false beliefs? Or can it be shown that unless preference is given to hypotheses that *appear* to be either warranted, true, or instances of knowledge, human beings will

---

*What shall we say about competing hypotheses (say in psychology) the acceptance of which would diversely affect the epistemic motivation of natural scientists? Is the more encouraging one to be preferred, other things being equal?

lose their motivation for scientific inquiry? In that case a certain kind of human motivation would supply the utilitarian link – the human desire, we might say, for instant epistemic gratification. These are a few of the interesting questions that come to mind once we acknowledge the fact that there is an irreducible distinction between two kinds of epistemic merit – intrinsic and instrumental.*

I am indebted to the following for questions and suggestions that have led to changes in the draft presented to the APA in December, 1980: Hannah Ginsborg, Bredo Johnsen, Hilary Putnam, Robert Shope, Ernest Sosa.