

Internalism and Externalism

"All knowledge is knowledge of someone; and ultimately no one can have any ground for his beliefs which does not lie within his own experience." C. I. Lewis¹

INTRODUCTION

We now turn to the dispute between those who would interpret epistemic justification "internally" and those who would interpret it "externally."²

The dispute concerns the proper analysis of the concept of epistemic justification; it presupposes, therefore, that the internalists and externalists share a common concept of justification—the one that distinguishes knowledge from true belief that is not knowledge.

We must be on guard, however, in interpreting contemporary literature that professes to be about "internalism" or "externalism." Some of those authors who profess to view knowledge and epistemic justification "externally" are not concerned with traditional theory of knowledge. That is to say, they are not concerned with the Socratic questions, "What can I know?" "How can I be sure that my beliefs are justified?" and "How can I

¹*An Analysis of Knowledge and Valuation* (La Salle: The Open Court Publishing Company, 1946), p. 236.

²For early statements of the distinction, see "The Internalist Conception of Justification," by Alvin Goldman, in *Midwestern Studies in Philosophy*, Vol. V (1980), pp. 27–51, and "Externalist Conceptions of Empirical Knowledge," by Laurence Bonjour, *ibid.*, pp. 53–74.

improve my present stock of beliefs?" Indeed many such philosophers are not concerned with the analysis of any ordinary concept of knowledge or of epistemic justification. Therefore their enterprise, whatever it may be, is not that of traditional theory of knowledge. My concern in what follows pertains only to the epistemological dispute, Is the concept of epistemic justification to be analysed internally or externally?

I will begin by saying what I understand by "internalism."

WHAT IS "INTERNALISM"?

The usual approach to the traditional questions of theory of knowledge is properly called "internal" or "internalistic." The internalist assumes that, merely by reflecting upon his own conscious state, he can formulate a set of epistemic principles that will enable him to find out, with respect to any possible belief he has, whether he is *justified* in having that belief. The epistemic principles that he formulates are principles that one may come upon and apply merely by sitting in one's armchair, so to speak, and without calling for any outside assistance. In a word, one need consider only one's own state of mind. But if we look at the matter, it seems clear that the approach to the questions of the traditional theory of knowledge can only be thus internalistic.

To be sure, we can assess the beliefs that *other* people have without examining *their* states of mind. And we can assess the beliefs that we ourselves have had at *other times* without examining the states of mind that *we* had at those other times. But these assessments although "external" in one sense, are "internal" in another.

Suppose we are considering the beliefs that some other persons had yesterday. After the fact, we may have information enabling us to assess their beliefs and to note just where they made their mistakes and where they did not. The principles we use need not be principles that were "internal" for them at the time that they had the beliefs in question. That is to say, the principles need not be principles that *they* could then have applied by reflecting upon their own states of mind. For the beliefs make use of information that is now available to *us* and was not then available to them. Hence they do not tell us anything about what they were then justified in believing about themselves. So far as they were then concerned the beliefs were "external"; they could not apply the principles merely by reflecting upon their states of mind. But if *we* are able to use the principles to appraise the beliefs of others, then the principles do presuppose something about what *we* are internally justified in believing about *them*.

According to this traditional conception of "internal" epistemic justification, there is no *logical* connection between epistemic justification and truth. A belief may be internally justified and yet be *false*. This consequence is not acceptable to the externalist. He feels that an adequate account of

epistemic justification should exhibit *some* logical connection between epistemic justification and *truth*.

In recent years there have been many proposals as to how epistemic justification might be explicated externally. But these suggestions, so far as I have been able to see, are of two sorts: either (1) they are empty or (2) they can be made to work only if they are supplemented by *internal* justification concepts. If this is true, then it has not yet been shown that internal concepts may be *replaced* by external ones.

I will consider, then, a number of externalistic attempts to explicate, "S is epistemically justified in believing p." I will suggest that some of them are empty (an "empty" explication being one that reduces justified belief to true belief). Then I will ask, with respect to those external explications that are not empty, whether they are adequate as they stand or whether they require supplementation by some epistemic concept that has not been shown to be externalistic.

THE NON-THEORY

I begin with a definition of external justification that is obviously unsatisfactory. I will use it to measure *other* possible definitions. For we may ask whether they tell us anything more than *it* does. We consider, then, theory (N)—"the non-theory":

(N) S is externally justified in believing p = Df p is true; and S is a thinking
subject □

The effect of this definition is to equate "external justification" with truth. Or, more exactly, the definition makes no distinction between the true beliefs that a person has and those beliefs that he is *justified* in having. I think it is fair to call this theory empty, since it does not contribute anything of significance to the theory of knowledge.

Can we, then, find a concept of "external" justification which does *not* thus reduce external justification to truth? Two types of external theory have been proposed—reliability theories and causal theories. And these may be combined into mixed reliability and causal theories.

I now turn to reliability theories.

RELIABILITY THEORIES OF EXTERNAL JUSTIFICATION

A common "reliability" definition of "external justification" is the following.³

³Compare Alvin Goldman: "Beliefs are justified if and only if they are produced by (relatively) reliable belief-forming processes." Op. cit., p. 47.

- (R1) S is externally justified in believing $p = \text{Df}$ The process by means of which S was led to believe p is reliable

One serious difficulty with the definition, as it stands, is that it does not allow us to say, of a person who does not believe p , that he is justified in believing p . But conceivably, by making judicious use of counterfactuals, one could repair the definition to provide for this possibility.

A more serious difficulty has to do with the interpretation of the expression "reliable process." If we take "process" in its broadest sense, then we may say that a *process* by means of which one is led to a belief is a series of *activities* that result in one's acquiring or retaining that belief. If we understand "process" this way and if "reliable process" means no more than, "process that is productive of true belief," then (R1) does not differ from (N)—that is to say, the present version of the reliability theory does not differ from the non-theory. For if the belief is true, then the process that led to it, however bizarre the process may have been, produced a belief that is true.

One may now want to say:

- (R2) S is externally justified in believing $p = \text{Df}$ The process by means of which S was led to believe p is a process which generally leads to true belief

Does this add anything to (N)? If S has acquired a true belief, then once again, no matter how bizarre the situation may be, he has followed *some* procedure which is such that following that procedure *always* leads to true belief.⁴ Let us consider this point in more detail.

If a person S has arrived at a true belief on a particular occasion, then S will have followed *some* procedure which was *unique* to that occasion. For example, S could have arrived at his belief by reading the tea leaves on a Friday afternoon twenty-seven minutes after having visited his uncle. If necessary, we may add further specifications—say, something about what S has just eaten or about the clothes that he is wearing. Since he has used this successful procedure *only on one occasion*, we may say:

- (e) S has arrived at the belief that p by means of a belief-forming process which is such that, whenever he arrives at a belief by means of that process, the belief he thus arrives at is true

If what we have said is correct, then *every* belief that S has arrived at will be one that has been arrived at by a unique process of the sort that (e) describes. Hence there is a process which is equivalent to the disjunction of

⁴See the discussion of this general question in Richard Feldman, "Reliability and Justification," *The Monist*. Vol. 68 (1985), pp. 159–174.

all those successful belief-forming processes and which has provided S with as many justified beliefs as he has true beliefs.

We have, then, a counter-example to the analysis set forth in R2. It may seem, at first consideration, that a simple repair will save the definition. To see that this is so, consider the following dialogue between the reliabilist (R) and the internalist (I):

(R) "You need only specify that the process not be a disjunctive process."

(I) "A *disjunction* is a type of sentence; but what is it for a *process* to be disjunctive?"

(R) "A process is disjunctive if it can be described using disjunctive sentences."

(I) "But *every* process can be described using disjunctive sentences; and therefore, if what you say is right, *every* process is disjunctive."

(R) "No. What I mean is that a disjunctive process is a process that can be described *only* by using disjunctive sentences."

(I) "But there is *no* process which can be described *only* by using disjunctive sentences . . ."⁵

The problem is that the following two propositions are true: (1) any disjunction of particular-procedures is such that, if we know enough about it, we can show it *also* to be a particular-procedure; and (2) any particular-procedure is such that, if we know enough about it, we can show it *also* to be a disjunction of particular-procedures. If you describe for me a procedure which you think is a disjunction of particular-procedures, I can add details which will entitle us to call it a particular-procedure, and if you describe for me a procedure which you think is a particular-procedure, I can add details which will entitle us to call it a disjunction of particular-procedures.

These observations are not intended to belittle the concept of a *reliable belief-forming process*. They are intended, rather, to belittle the suggestion that epistemic justification can be defined merely by reference to such processes. Obviously one should try to *know* what belief-forming processes one is following and one should try to *find out* which of those processes are

⁵In *Epistemology and Cognition* (Cambridge, MA: Harvard University Press, 1985), Alvin Goldman suggests other moves that the reliabilist might make to repair R2. Thus the somewhat bizarre example of the tea leaf reader could be avoided if we restricted our description of belief-forming processes to *organic processes* within the body of the believer (see p. 50). But here, too, there will be a unique bodily process for every belief-acquisition. We can all now truly say: "I never was in exactly *this* bodily state before and I never will be in it again." Should we add, then, that the processes be processes that are *relevant* to the acquisition or retention of belief? This move, of course, transfers the problem to that of finding a suitable analysis of "relevant."

reliable; then one should try as far as possible to follow them. But this is to say that we should be concerned to follow those processes which are such that we are *justified* in believing them to be *reliable*.

Consider, now, the following definition:

- (R3) S is externally justified in believing $p = Df$ The process by means of which S was led to believe p is one which is such that it is *evident* to S that that process generally leads to true belief

Since "evident" expresses one of the internalist's epistemic concepts and since no externalistic explication of the concept of *being evident* is at hand, we may say this of (R3): it is an analysis of external justification which *combines* internal and external justification concepts. We could replace "evident" in (R3) by "knows" and say that the process is one which is such that S *knows* that it generally leads to true belief. If no externalistic explication of knowledge is added, then, once again, we have a definition that *combines* internal and external concepts.

Another possibility is to construe a *reliable* process as a process which is *probably* such that it leads to truth.⁶ Then we might have:

- (R4) S is externally justified in believing $p = Df$ The process by means of which S was led to believe p is a process which is *probably* such as to lead to true belief.

We have discussed in detail the *epistemic* concept of probability: the concept we have discussed is externalistic. Principle R4 has a certain plausibility if we take the word "probably" in this epistemic sense. But how is the principle to be applied if we take "probably" in its *external* sense? The difficulty may be seen by contrasting two of the principal uses of "probable"—the *statistical* use and the *relational* use.⁷

Statements in which "probable" is taken merely *statistically* are reword-

⁶Laurence Bonjour writes: "If finding epistemically justified beliefs did not substantially increase the likelihood of finding true ones, then epistemic justification would be irrelevant to our main cognitive goal and of dubious worth"; *The Structure of Empirical Knowledge* (Cambridge, MA: Harvard University Press, 1985), p. 8. Compare Ernest Sosa: "Faculty F is *more reliable* than faculty F' if the *likelihood* with which F would enable one to discriminate truth from falsehood in f (F) is higher than the likelihood with which F' would enable one to make such discrimination in f (F')." I have italicized "likelihood." (It should be noted that Sosa here speaks of faculties instead of belief-yielding processes.) See Ernest Sosa, "Knowledge and Intellectual Virtue," *The Monist*, Vol. 68 (1985), pp. 226–247; the quotation is on p. 238.

⁷See Rudolf Carnap, *Logical Foundations of Probability* (Chicago, IL: The University of Chicago Press, 1950), p. 300ff. Carnap speaks of statistical probability statements as statements about "probability²": and of relational probability statements as statements about "probability¹". He notes that "under certain conditions, probability¹ may be regarded as an estimate of probability²" (p. 300). But in the theory of knowledge, statements of *relational probability* are not concerned merely with estimates of relative frequencies; a typical statement of relational probability would be: "Thinking that one remembers p tends to make p probable."

ings of statements about statistical frequencies; they state what proportion of the members of one class are also members of another class. For example, "The probability that any given A is a B is n ," might be interpreted as telling us, " n percent of the members of the class of A's are also members of the class of B's." (Statisticians make use of interpretations that are considerably more complex, but the added complexity does not affect the points that are here at issue.) What do statements about statistical frequencies tell us about the justification of belief?

It would not be helpful to say merely that a belief is justified provided only that it is a member of a set of beliefs most of whose members are true. For this would have the consequence that every belief is justified. Moreover, it would not help us find a plausible interpretation of R4. Our subject S had followed a belief-forming process which was such that *all* the beliefs that he arrived at by using that process were true. Hence the statistical probability of that process yielding a true belief would be a probability of 1. Taking explication (R4) this way, we do not progress beyond the original explication (N).

Or could the reliabilist take probability in its *relational* sense—in the sense of, "h is probable in relation to e"? We have seen how the concept of the *evident* may be used in explicating the application of this relational concept to the beliefs of a particular subject. But if the reliabilist does not make use of the concept of the evident how will *he* apply the relational concept?

Perhaps he will say this, "A proposition h is probable *for* a particular subject S, provided only: there is a *true* proposition e which is such that h is probable in relation to e and S accepts e." May we say of the proposition p, which S had arrived at as a result of his bizarre belief-forming process, that there is a *true* proposition which is such that p is more probable than not in relation to that proposition? One such proposition is our earlier proposition (e):

- (e) S has arrived at the belief that p by means of a belief-forming process which is such that, whenever he arrives at a belief by means of that process, the belief he thus arrives at is true

And so we have not found any purely external sense of "probability" in terms of which we can interpret the reliabilist's R4.

There are, of course, other statistical and relational interpretations of "probable," but, so far as I have been able to see, none of them is of any help to the externalist.⁸

⁸As we have seen, some would interpret the relational sense of "probability" without appeal to the concept of evidence and would say that a proposition is probable for a given person provided only that the proposition is probable in relation to what that person happens to believe. It is difficult to see how *this* way of construing probability would provide us with an account of epistemic justification.

CAUSAL THEORIES OF EXTERNAL JUSTIFICATION

Some have hoped to establish a connection between justification and truth by defining justification by reference to causation. Consider, for example, those true propositions which are such that their *being true* is what causes us to believe that they are true. Could it be that these are the propositions we are “externally” justified in believing? At best, this suggestion gives us a very restricted account of epistemic justification. For it is not applicable as it stands to propositions about the future. And it is doubtful whether it would be applicable to propositions that are logically true. Are there, however, some propositions that may be said to be justified in this way?

The locution, “A causes B,” may be taken in two quite different ways—(1) as telling us that A is *the cause of B* or (2) as telling us that A *contributes causally* to B (that A is one of the *causal factors* that lead to B). We have, then, two causal definitions to consider.

The first is this:

- (C1) S is externally justified in believing $p = Df$ S believes p ; and p 's being true is *the cause of S's believing p*

The phrase “the cause” is certainly one that is in common use; indeed it is suggested by the familiar propositional connective, “because.” Thus many people like to think that, of the various events that contribute causally to a given event, there is just one of them may properly be singled out as *the cause* of that event. Such a view is especially tempting when we are looking for a scapegoat.⁹ But, as we know from the study of the nature of causation, the expression, “A is *the cause of B*,” is one that is applicable only in very restricted circumstances and is not likely to be of use in connection with the present problem. If p , for example, is the proposition that there are mountains on the other side of the moon, then it is doubtful whether one could pick out *any* situation in which p 's being true could be said to be *the cause* of anyone's belief that p . That event which is p 's being true is just one of many factors which, working together, contribute causally to the belief that p .

What if we were to define, “A is *the cause of B*,” by saying, “Of those events that contribute causally to E, A is the sole change that immediately preceded the occurrence of E”?¹⁰ If we take “the cause” this way, then the

⁹We might say of the expression “the cause” what William James said of “cause”—namely that it is “an altar to an unknown God, an empty pedestal still marking the place of a hoped for statue.” William James, *The Principles of Psychology* (New York: Henry Holt and Company, 1893), Vol. II, p. 671.

¹⁰C. J. Ducasse proposed that “the cause of a change K” is that change which “alone occurred in the immediate environment of K immediately before”; *Truth, Knowledge and Causation* (London: Routledge & Kegan Paul, 1968), p. 4. Ducasse's definition, unlike the one proposed above, did not make use of the concept of causal contribution (causal factor).

cause of the acquisition of a belief might be some other psychological event (the occurrence, say, of a certain thought) or it might be some neuro-physiological event. Application of (C1), therefore, would be restricted to those beliefs which are about such psychological or neuro-physiological events.

Does the causal theory fare better if we replace "is the cause of" by "causally contributes to"? Then we would have:

(C2) S is externally justified in believing p = Df S believes p; and p's being true contributes causally to S's believing p

Now the definition is subject to Rube Goldberg counter-examples. Consider a person who is working in the garden and who suddenly becomes tired. His fatigue leads him to go inside and read the newspaper. He reads that some of the people who suffer from a certain internal disorder have red hair. Since *he* has red hair and is also a hypochondriac, he concludes, "I've got that disorder!" If, now, his having that disorder was one of the many factors that contributed causally to his fatigue, then we may say that, according to (C2), he is externally justified in believing that he has that disorder. *This* concept of justification is not likely to be of use in investigating the theory of knowledge.

Could one overcome such difficulties by specifying a type of causation that is not transitive? *Direct causation* is not transitive. (Roughly, "A is a direct causal contributor to B, if and only if: A contributes causally to B, and A does not contribute causally to anything that contributes causally to B.") The direct contributor to a belief attribution would then presumably be either another psychological state or an internal physiological state. This move, then, has the same difficulties as the version of "the cause" move considered above.

Our example above may suggest that the subject S should be *aware of* the causal role that is played by *p* in the formation of his belief. And so one might suggest:

(C3) S is externally justified in believing p = Df S believes p; p's being true contributes causally to S's believing p; and it is *evident* to S that p's being true contributes causally to his belief that p

This proposal is like R3 above: it *combines* internal and external justification concepts.¹¹

¹¹This type of theory is suggested by Marshall Swain who proposes a causal theory that makes use of such internalistic expressions as the following: "S's evidence" and "renders evident." See Marshall Swain, "Knowledge, Causality and Justification," in Pappas and Swain, *op. cit.*, pp. 87-99.

MIXED THEORIES

The reliability and causal theories that we have considered may be combined in various ways.¹² We need consider only two possibilities:

- (M1) S is externally justified in believing p = Df S believes that p; and *the cause* of S's believing that p is that S follows a belief-forming process that *generally leads to true belief*

This combines R2 and C1 and obviously has the difficulties of each.

- (M2) S is externally justified in believing p = Df S believes that p; and one of the facts that *contributes causally* to his believing p is the fact that he followed a belief-forming process which, *more probably than not*, yields true belief

This combines R4 and C2 and obviously has the difficulties of each.

The "externalistic" explications of epistemic justification that we have considered are all such that either they are empty or they make use of internal concepts. It would seem, therefore, that there is no indication that externalistic justification concepts may *replace* internal concepts.

¹²"Reliabilism is the view that a belief is epistemically justified if and only if it is *produced* or *sustained* by a cognitive process that reliably yields truth and avoids error." Sosa, op. cit., p. 239. I have italicized the causal expressions "produced" and "sustained."