

## II. VARIETIES OF PRIVILEGED ACCESS

WILLIAM ALSTON

IT is a very common, though by no means uncontested, view that the kind of knowledge a person has of his own mental (psychological) states, such as thoughts and feelings, is in principle not only fundamentally different from but also superior to the knowledge of his thoughts and feelings that is available to anyone else. Following an established usage, we may express this view by saying that a person has "privileged access" to his own mental states. It is obvious that this thesis will vary in content with variations in the specific mode of superiority imputed. Nevertheless, discussions of privileged access, both pro and con, have never been sufficiently alive to these variations or to their significance.

The central task of this paper is the exhibition and interrelation of the most important of the ways in which one's access to one's own mental states has been, or might be, thought to be privileged. In addition I shall show, though only sketchily, how failure to be alive to the full range of possibilities has vitiated some prominent discussions of the topic.

### I

First, a couple of preliminary points. Privileged access claims vary not only with variations in the mode of epistemic superiority imputed, but also with variations in the category of "mental states" with respect to which the claim is made. Many philosophers have advanced privileged access as a *criterion* for the mental or the psychological; they have held that a state of a person is mental (psychological) if and only if that person's knowl-

edge that he has the state is in principle superior, in some specified way, to the knowledge of that fact that is available to anyone else.<sup>1</sup> Others have made distinctions within this class of states and have asserted their favored form of privileged access of some sub-class thereof. Thus it is not uncommon to hold that one cannot be mistaken with respect to what may be called "phenomenal states," i.e., present contents of consciousness, such as sensations, images, feelings, and thoughts, but not to assert infallibility with respect to what may be called "dispositional states," such as beliefs, desires, and attitudes.<sup>2</sup> However the most common procedure is simply to work with particular examples, e.g., sensations or, more specifically, pains, and not even attempt to make clear the several classes of states of which privileged access is being asserted.<sup>3</sup> In this paper I mention this dimension of variation only to set it aside. My sole concern will be to distinguish and compare various types of epistemic superiority; I shall not also be concerned to distinguish and compare various classes of entities with respect to which one or another of these has been asserted. Hence for our purposes we can just work with the rather loose rubric, "mental state," remembering that if anyone is to put forward a privileged access thesis, he should be more specific as to the range of states involved.

We can hardly avoid taking note, however briefly, of those philosophers who would make short shrift of our entire problem by dismissing it, on the grounds that it makes no sense to speak of a person *knowing* that he has, e.g., a certain sensation. (See L. Wittgenstein, *Philosophical Investigations*, Pt. I, para. 246.) If that is the case, there is

<sup>1</sup> See, e.g., G. E. Moore, "The Subject Matter of Psychology," *Proceedings of the Aristotelian Society*, vol. 10 (1909-1910), pp. 36-62, reprinted in G. N. A. Vesey (ed.), *Body and Mind* (London 1964); and F. Bretano, selection from *Psychology from an Empirical Standpoint*, in *ibid*.

<sup>2</sup> See, e.g., J. Shaffer, "Persons and Their Bodies," *The Philosophical Review*, vol. 75 (1966), pp. 59-77.

<sup>3</sup> Thus Norman Malcolm, in "Direct Perception" (in *Knowledge and Certainty*, Englewood Cliffs, N.J., 1963) restricts his discussion to after-images; while in "The Privacy of Experience" (in A. Stroll, ed., *Epistemology, New Essays in the Theory of Knowledge*, New York, 1967), he specifically discusses pain, and sometimes more generally "sensations." Presumably Malcolm supposes that the things he says about after-images (pains) have a wider scope of application, but he does not make explicit just what he takes this to be. Again, most of A. J. Ayer's discussion in his essay "Privacy" (in *The Concept of a Person and Other Essays*, New York, 1963) is in terms of "thoughts and feelings," but he makes no attempt to say exactly how far he means his remarks to extend.

no problem as to whether one's own knowledge of his own sensations is in some way necessarily superior to that available to any other person. I cannot really go into the issue in this paper, but it may not be out of place to explain briefly why it seems to me that any argument for this conclusion must be defective in *some* way. Clearly someone else can be in doubt as to whether I am in a given mental state, e.g., whether I am thinking about tomorrow's lecture, whether I am worrying about my job prospects, whether I feel elated. That is, he may not know how to answer a certain question, "Does he (Alston) feel elated?" But it seems that normally I *would* be in a position to answer that question, the *same* question to which he does not know the answer. But how can we understand my being in that position without supposing that I know something he doesn't, e.g., that I do feel elated? Thus it seems to be as undeniable as anything could be that persons normally do know what mental states they are in at a given moment, and that no argument designed to show that this is false or meaningless can be sound.

## II

I shall begin by extracting a number of possible modes of privileged access from a rather wide sampling of the literature. We may begin with the following.

Am I not that being who now doubts nearly everything, who nevertheless understands certain things, who affirms that one only is true, who denies all the others, who desires to know more, is averse from being deceived, who imagines many things, sometimes indeed despite his will, and who perceives many likewise, as by the intervention of the bodily organs? Is there nothing in all this which is as true as it is certain that I exist, even though I should always sleep and though he who has given me being employed all his ingenuity in deceiving me? . . . Finally, I am the same who feels, that is to say, who perceives certain things, as by the organs of sense since in truth I see light. I hear noise, I feel heat. But it will be said that these phenomena are false and that I am dreaming. Let it be so; still it is at least quite certain that it seems to me that I see light, that I hear noise, and that I feel heat. That cannot be false; . . .

— R. Descartes, *Meditations*, II

. . . for a man cannot conceive himself capable of a greater certainty than to know that any idea in his mind is such as he perceives it to be; and that two ideas, wherein he perceives a difference, are different and are not precisely the same.

— J. Locke, *Essay Concerning Human Understanding*, IV, 2

For since all actions and sensations of the mind are known to us by consciousness, they must necessarily appear in every particular what they are, and be what they appear. Everything that enters the mind, being in *reality* as the perception, tis impossible anything should to *feeling* appear different. This were to suppose that even where we are most intimately conscious, we might be mistaken.

— D. Hume, *Treatise of Human Nature*, I, iv, 2

The facts of consciousness are to be considered in two points of view; either as evidencing their own ideal or phaenomenal existence, or as evidencing the objective existence of something else beyond them. A belief in the former is not identical with a belief in the latter. The one cannot, the other may possibly be refused. . . . Now the reality of this, as a subjective datum—as an ideal phaenomenon, it is absolutely impossible to doubt without doubting the existence of consciousness, for consciousness is itself this fact; and to doubt the existence of consciousness is absolutely impossible; for as such a doubt could not exist, except in and through consciousness, it would, consequently, annihilate itself.

— Sir W. Hamilton, *Lectures on Metaphysics*, XV, p. 188

It is a further general characteristic of all mental phenomena that they are perceived only in inner consciousness. . . . One could believe that such a definition says little, since it would seem more natural to take the opposite course, defining the act by reference to its object, and so defining inner perception of mental phenomena. But inner perception has still another characteristic, apart from the special nature of its object, which distinguishes it: namely, that immediate, infallible self-evidence, which pertains to it alone among all the cases in which we know objects of experience. Thus, if we say that mental phenomena are those which are grasped by means of inner perception, we have accordingly said that their perception is immediately evident.

— F. Brentano, *Psychology from an Empirical Standpoint*, selection in G. N. A. Vesey, ed., *op. cit.*, p. 151.

Subtract in what we say that we see, or hear, or otherwise learn from direct experience, *all that conceivably could be mistaken*; the remainder is the given content of the experience inducing this belief. . . . Apprehensions of the given which such expressive statements formulate, are not judgments, and they are not here classed as knowledge, because they are not subject to any possible error. Statement of such apprehension is, however, true or false; there could be no doubt about the presented content of experience as such at the time when it is given, but it would be possible to tell a lie about it.

— C. I. Lewis, *An Analysis of Knowledge and Valuation*, La Salle, Ill.: Open Court, 1946, pp. 182–183

Some Philosophers . . . have thought it possible to find a class of statements which would be both

genuinely informative and at the same time logically immune from doubt. . . . The statements usually chosen for this role . . . characterize some present state of the speaker, or some present content of his experience. I cannot, so it is maintained, be in any doubt or in any way mistaken about the fact. I cannot be unsure whether I feel a headache, nor can I think that I feel a headache when I do not.<sup>4</sup>

– A. J. Ayer, *The Problem of Knowledge*, London: Macmillan & Co., 1956, p. 55

Besides what is logically certain there are a number of immediately known propositions which we can regard as absolutely certain although there would be no self-contradiction in denying them. In this class I put more specific propositions based on introspection. I cannot see any self-contradiction in supposing that I might make mistakes in introspection; and there is therefore no *logical* absurdity in supposing that I might be mistaken now when I judge that I feel warm or that I have a visual presentation of a table. But I still cannot help being absolutely certain of the truth of these propositions and I do not think that I ought to be otherwise. . . . As we have seen, it is however hardly possible to claim this absolute certainty for judgments about physical objects, and, as we shall see, there are similar difficulties in claiming it for judgments about minds other than one's own.

– A. C. Ewing, *The Fundamental Questions of Philosophy*, ch. V

I think the facts that give rise to the illusion of privacy would be the following: (a) you can be *in doubt* as to whether I am in pain, but I cannot; (b) you can *find out* whether I am in pain, but I cannot; and (c) you can be *mistaken* as to whether I am in pain, but I cannot.

– N. Malcolm, "The Privacy of Experience," in A. Stroll, ed., *op. cit.*, p. 146

But there is also a sense in which a person's report that he sees an after-image *cannot* be mistaken; and it is this sense that I intend when I say that his report is "incurrable."

– N. Malcolm, "Direct Perception," in *Knowledge and Certainty*, p. 85

Among the incurrable statements are statements about "private" experiences and mental events, e.g., pain statements, statements about mental images, reports of thoughts, and so on. These are incurrable in the sense that if a person sincerely asserts such a statement it does not make sense to suppose, and nothing could be accepted as showing, that he is mistaken, i.e., that what he says is false.

– S. Shoemaker, *Self-Knowledge and Self-Identity*, Ithaca, N.Y.: Cornell U. Press, 1963, pp. 215–216

All of these quotations represent one's epistemic position *vis-à-vis* one's own mental states (or some subclass thereof) as highly favorable in some way or other. In most of the passages quoted there is no contrast explicitly drawn with the epistemic position of other persons, but such a contrast is implicit in what is said. None of these philosophers would suppose that other persons have the kind of cognitive access to my mental states which they impute to me; hence by being in this kind of position one enjoys a kind of special epistemic privilege.

How many distinguishable types of favorable epistemic position are involved in these passages? One type that is clearly imputed in several of the quotations is the impossibility of mistake. Thus one's judgments or beliefs about his own mental states "cannot be false" (Descartes), "are not subject to any possible error" (Lewis), "cannot . . . be . . . in any way mistaken" (Ayer), "it does not make sense to suppose that he is mistaken" (Shoemaker). A great many terms have been used for this kind of epistemic privilege. I prefer "infallibility."

There is also much talk in these passages about immunity from doubt (Descartes, Hamilton, Lewis, Ayer, Ewing, Malcolm). But we can distinguish several different indubitability claims, each of which can be attributed to one or more of our authors. First there is the claim that it is impossible to *entertain a doubt* as to the truth of a proposition attributing a current mental state to oneself. I am *incapable of being in doubt* as to whether I am now thinking about my lecture for tomorrow, or whether there is now an image of my boyhood home before my mind's eye. This impossibility might be logical, or it might be nomological (based, e.g., on psychological laws). Malcolm in the first quotation is clearly asserting the former, for he asserts the three points as facts about the "grammar" of the word "pain"; because of the way we use the word, no sense can be attached to speaking of a person having a doubt as to whether he is in pain. On the other hand Lewis might be plausibly interpreted as claiming that it is a psychological impossibility for one to doubt whether he is currently in some conscious state.

However, we can also discern a quite different concept of indubitability at work in these authors. This is a normative rather than a factual concept—not the impossibility of being, in fact, in a psychological state of doubt, but rather the impossibility of having any grounds for doubt, the impossibility

<sup>4</sup> It should be noted that this passage sets forth a view that Ayer is examining rather than propounding.

of a rational doubt. In our quoted material this comes out most clearly in the passage from Hamilton. What he is arguing there, in the spirit of Descartes, is not so much that there are psychological bars to the formation of a doubt, but rather that such a doubt would necessarily lack any foundation, since it presupposes that which is called into question, viz., the fact of consciousness.<sup>5</sup> Again, if, as we shall argue later, we can take Ewing to be using "certainty" as equivalent to "indubitability," he clearly distinguishes our two main senses of indubitability and asserts both. "But I still cannot help being absolutely certain of the truth of these propositions and I do not think that I ought to be otherwise." That is, I find it psychologically impossible to have any doubt of their truth, and I am justified in this incapacity, since there could be no grounds for any doubt.

Thus we have distinguished three forms of indubitability: logical impossibility of entertaining a doubt, psychological impossibility of entertaining a doubt, impossibility of there being any grounds for doubt. Although for any of them it is worth considering whether propositions about one's own current mental states are indubitable in that sense, still it is only the third that constitutes a distinctively *epistemic* privilege. If I am so related to a certain group of propositions that whenever I believe one of those propositions to be true, there can be no grounds for doubt that it is true, then I am in a very favorable position to obtain knowledge in this sphere; for, unlike the usual situation, whatever I believe, no one can have any justification for refusing to accept my belief as true. But that means that I have every right to accept the proposition; so that in this sphere, each and every one of my beliefs will automatically count as knowledge. This is certainly to be in a highly favorable epistemic position. On the other hand, the mere fact that I find it psychologically impossible to doubt the truth of any such proposition does not in itself confer any cognitive superiority. We can think of many cases where people are unable to entertain doubts about certain matters, and where we regard this as a liability rather than an asset. Very small children are often unable to imagine that what their parents say is mistaken, and religious fanatics are sometimes psychologically unable to doubt the tenets of their sect. We do not take such people to be thereby in a better position for acquiring knowledge; quite the contrary, we suppose this critical

incapacity to be hampering them in the cognitive enterprise. To be sure, when the second sort of indubitability is imputed to propositions about one's own mental states, it is supposed that this holds for all men as such, and it may be thought that this renders inappropriate an epithet such as "lack of critical faculty" which one might suppose to be applicable only when the disability in question is peculiar to certain stages of development, types of personality, or kinds of social groups. Nevertheless, if we suppose that a universally shared psychological inability to doubt confers some advantage in the acquisition of knowledge, it is only because we think that this *psychological* inability is conjoined with, and perhaps is a reflection of, indubitability in the normative sense, the impossibility of any *grounds* for doubt. If it *should* be the case that the psychological impossibility of doubting the truth of one's beliefs about one's own current mental states is due to an ingrained weakness in the human critical apparatus, or to an irresistible partiality to one's own case, then this inability would *not* indicate any first person epistemic advantage in these matters.

We shall have to make the same judgment concerning Malcolm's thesis of the logical impossibility of entertaining a doubt. Suppose we grant that the meanings we attach to our conscious state terms are such that it makes no sense to suppose that a given person is in doubt as to whether he is currently in a certain kind of conscious state. That would be a noteworthy feature of our conceptual scheme, but we still have to ask whether or not it is well founded. Unless we accept normative indubitability, or some other principle according to which a person is in a particularly favorable position to discriminate true from false propositions concerning his present conscious states, then we will have to conclude that the features of our "logical grammar" to which Malcolm alludes are ill-advised; and that the fact that this "logical grammar" is as it is does nothing to show that persons are in a specially favorable epistemic position *vis-à-vis* their own current conscious states.

Thus I conclude that normative indubitability is the only variety that clearly constitutes a cognitive advantage. We shall henceforth restrict the term "indubitability" to that variety.

We might think of indubitability as a weaker version of infallibility. To be infallible *vis-à-vis*

<sup>5</sup> The merits of Hamilton's argument, and any other *substantive* question concerning privileged access, are not within the jurisdiction of this article.



one's present conscious states is to be in such a position that no belief one has to the effect that one is in such a state *can* be mistaken. Whereas an indubitability thesis does not commit one to the impossibility of mistakes, but to the weaker claim that no one could have grounds for questioning the accuracy of one's belief. There is a still weaker derivative of infallibility that can be found in the literature, though more rarely. It is set out clearly in the following passage from A. J. Ayer's British Academy lecture on "Privacy."

If this is correct, it provides us with a satisfactory model for the logic of the statements that a person may make about his present thoughts and feelings. He may not be infallible, but still his word is sovereign. The logic of these statements that a person makes about himself is such that if others were to contradict him we should not be entitled to say that they were right so long as he honestly maintained his stand against them.<sup>6</sup>

What Ayer is saying here is that it is impossible that anyone else should show that I am mistaken in what I say (believe) about my present thoughts and feelings. This is an inherent impossibility, for the "logic of these statements" requires us to give the person in question the last word. We may term this kind of epistemic position "in corrigibility."<sup>7</sup> In corrigibility is weaker than indubitability, for whereas the latter rules out the possibility of any grounds for doubt, however weak, the former only rules out someone else's having grounds for the contradictory that are so strong as to be sufficient to *show* that I was mistaken.<sup>8</sup>

"Certainty" is another term that figures prominently in our quotations. A person's judgments concerning his own mental states are said to exhibit the highest degree of certainty; one can be *absolutely* certain about such matters (Descartes, Locke, Ewing). How are we to interpret these claims vis-a-vis the others we have been considering? Here too we may distinguish factual and normative senses. Being certain of something may be construed as a matter of feeling assurance, feel-

ing confident that one is correct; this is presumably the reverse side of the (*de facto*) absence of doubt. To feel complete confidence that one is correct is to entertain no doubt about the matter. And a psychological (logical) impossibility of the entertaining of any doubt would be the same thing as a psychological (logical) necessity of feeling completely assured that one is correct. Thus this kind of certainty comes under the scope of the arguments just given to dismiss the corresponding forms of indubitability from further consideration.

However there is also a normative concept of certainty, a concept employed by Ewing when after saying "I still cannot help being absolutely certain of the truth of these propositions" he adds, "... and I do not think that I ought to be otherwise." To be certain in this sense is to be justified in feeling complete assurance. How is this normative concept related to the modes of epistemic superiority already distinguished? It seems impossible to make a general identification of normative certainty with any of the other modes. To be justified in feeling complete assurance that *S* is to have a very strong warrant for one's belief that *S*. But views may differ as to just how strong a warrant is required: the strongest conceivable, the strongest one could reasonably ask for in the subject matter under consideration, and so on. Thus the general concept of normative certainty is really a sort of family or continuum of concepts, differing as to the chosen locus along the dimension of strength of warrants for belief. Whereas our other modes of epistemic superiority are not subject to variations in degree; they are absolute concepts. If one's belief is indubitable, *no* doubt can have any basis: if one is infallible, one's belief must be *wholly* correct; and so on.<sup>9</sup>

But although we cannot make any general identification of the concept of normative certainty with the other modes of epistemic superiority we have distinguished, still I think that the degree of certainty typically ascribed to one's beliefs about one's own mental states amounts either to infallibility or

<sup>6</sup> A. J. Ayer, "Privacy," in *op. cit.*, p. 73.

<sup>7</sup> The use of this term presents the usual chaotic picture. We have seen Malcolm and Shoemaker using it to mean infallibility. Thomas Nagel in his essay, "Physicalism," (*The Philosophical Review*, vol. 74 [1965], p. 344) uses it to mean what I shall next be distinguishing as "omniscience." I believe that the present usage is a more apt one.

<sup>8</sup> It may be contended, e.g., by partisans of the "private language argument," that there is no significant difference between an impossibility of anyone else's showing that I am mistaken and an impossibility of my being mistaken. I am unable to go into those issues in this paper.

<sup>9</sup> One could, of course, construct degree-concept derivatives of these absolute concepts. Thus one could distinguish various degrees of immunity to rational doubt, depending on what kinds of doubts are excluded, how strong or weak the grounds would have to be, and so on. However, so far as I know, such degree concepts are not in fact employed in connection with the present topic.

to indubitability. Sometimes it is claimed (Descartes, Locke) that such beliefs enjoy the highest conceivable certainty; in that case one is in effect ascribing infallibility, for the highest warrant one could conceive for a belief is one which would render the falsity of the belief *impossible*. In other cases something weaker is being claimed; thus Ewing conjoins his assertion of certainty with the admission that "there is no self-contradiction in supposing that I might make mistakes in introspection." However, in such cases it is plausible to suppose that a warrant strong enough to exclude all grounds for doubt is being imputed, and hence that what is being ascribed is indubitability. Thus I do not feel that we need "certainty" as a separate item in our list.

We have still not exhausted the conceptual riches of our initial list of quotations. Going back to the passage from Hume, we note that he not only says of "actions and sensations of the mind" that they "must necessarily . . . be what they appear," which is infallibility, but that they "must necessarily appear in every particular what they are." In other words, it is not only that every belief or judgment which I form about my present mental states must be correct; it is also necessary that every feature of those states must find representation in those (necessarily correct) beliefs. *Ignorance* as well as *error* is excluded. Let us use the term "omniscience" for the logical impossibility of ignorance concerning a certain subject-matter. Although the Hume quotation contains the only omniscience claim in our original list, we can find other passages in which it is asserted that one is omniscient *vis-à-vis* his own mental states.

It requires only to be stated to be admitted, that when I know, I must know that I know,—when I feel, I must know that I feel,—when I desire, I must know that I desire. The knowledge, the feeling, the desire, are possible only under the condition of being known, and being known by me.<sup>10</sup>

Thinking and perceiving are essentially conscious processes, which means that they cannot be said to occur unless the person to whom they are ascribed knows that they occur.<sup>11</sup>

We can better represent and interrelate the modes of epistemic privilege we have distinguished, and will be distinguishing, if we have a standard formula for favorable epistemic positions, a formula containing blanks such that when these blanks

are filled in differently we get specifications of different modes. One might at first suppose that our formula could simply be: *X's knowledge of \_\_\_\_\_ is \_\_\_\_\_*, where the first blank is filled with a specification of the subject-matter, and the second blank with a specification of a particular mode of cognitive superiority—infallibility, omniscience, or whatever. However this will not work, since it is not in general true that the modes we are distinguishing are features of pieces of knowledge, features which a given piece of knowledge might or might not have. This is particularly clear with respect to infallibility and omniscience. We cannot first ascertain that *P* knows that *S*, and then go on to ask whether that bit of *P's* knowledge is or is not infallible or omniscient. The reason is somewhat different in the two cases. Infallibility in the sense of *cannot be mistaken* is a feature necessarily possessed by every piece of *knowledge* in a strong sense of "knowledge." That is, it would not be correct to attribute knowledge that *S* to *P* unless *P's* supposition that *S* were correct. That is part of what we mean by "know." If I do not feel elated now, then that is enough to (logically) rule out the possibility that I, or anyone else, know that I am elated now. Thus infallibility does not constitute a feature that distinguishes one kind of *knowledge* from another. With omniscience (in a certain area) on the other hand, the point is that this is a feature of one's position with respect to the possession or non-possession of knowledge (of certain matters), rather than a feature of any particular instance of such knowledge; it is a matter of what kinds of knowledge one (necessarily) has, rather than a matter of the character of that knowledge once obtained.

But although infallibility and omniscience are not characteristics that (may) attach to some pieces of knowledge and not to others, they clearly have something to do with knowledge. They are, in some way, features of one's epistemic position, powers, or status, *vis-à-vis* some domain of knowledge. Perhaps we can find an illuminating way of representing these modes of privileged access if we dig into the structure of the concept of knowledge, rather than just using it in an unanalyzed form. For our purposes we can work with the following familiar tripartite analysis of "*P* knows that *S*." The analysans consists of a conjunction of the following:

<sup>10</sup> Sir William Hamilton, *op. cit.*, Lect. XI, p. 133.

<sup>11</sup> D. Locke, *Myself and Others*, ch. II, p. 17.

- A. *P* believes that *S*.
- B. *P* is justified in believing that *S*.
- C. It is the case that *S*.<sup>12</sup>

As for infallibility, although a piece of knowledge is not the sort of thing that may or may not be capable of error, there is, according to the above analysis, a constituent of *P*'s knowledge that *S* which may or may not be capable of error, viz., *P*'s belief that *S*.<sup>13</sup> Thus one can be said to be infallible *vis-à-vis* a certain subject matter provided one cannot be mistaken in any beliefs he forms concerning that subject matter.

A person enjoys infallibility<sup>14</sup> *vis-à-vis* his own mental states = *df.* It is logically impossible that a belief of his about his own mental states should be mistaken.

Now if one is so situated relative to a given belief, he is amply, indeed maximally, *justified* in holding that belief. For one could hardly have a stronger (epistemic) justification for holding a certain belief than the logical impossibility of the belief's being mistaken. Hence where the mere possession of the belief logically guarantees truth, it equally guarantees the belief's being justified, i.e., it guarantees the satisfaction of both the other two conditions for knowledge. Hence we can just as well state our definition as follows (generalizing now over subject-matters, so as not to restrict the general concept of infallibility to the topic of one's own mental states):

- (DI) *P* (a person) enjoys infallibility with respect to a type of proposition, *R* = *df.* For any proposition, *S*, of type *R*, it is logically impossible that *P* should believe that *S*, without knowing that *S*. (Condition A. for *P*'s knowing that *S* logically implies conditions B. and C.)

The philosophers who shy away from speaking of a person's *knowing* that he has certain thoughts and sensations will probably be even more leary of

speaking of a person's *believing* that he has a certain thought or sensation. And it must be admitted that one does not ordinarily speak in this vein. But, so far as I can see, this is simply because we ordinarily use the word "belief" in such a way that it contrasts with knowledge, as in the following dialogue:

What was that noise in the kitchen?

I believe that the tap was leaking.

You *believe* it was leaking! Couldn't you see whether it was or not?

In this paper, as quite frequently in philosophy, we are using the word in a wider sense. This sense can be indicated by making it explicit that a sufficient condition for *P*'s believing that *S* is that *P* would have a tendency to assert that *S* if he were asked whether it were the case that *S*, if he understood the question, and if he were disposed to be sincere. In this wider sense one often believes that he has certain thoughts and feelings. At this point the Wittgensteinian will, no doubt, cavil at the idea that one can correctly be said to *assert* that he has a certain feeling, but I cannot pursue the controversy further in this paper.

Omniscience can be given a parallel formulation as follows:

- (D2) *P* enjoys omniscience *vis-à-vis* a type of proposition *R* = *df.* For any true proposition, *S*, of type *R*, it is logically impossible that *P* should not know that *S*. (Condition C. for *P*'s knowing that *S* logically implies conditions A. and B.)<sup>15</sup>

Thus this familiar analysis of knowledge permits us to give a neat presentation of the infallibility-omniscience distinction. They differ just as to which of the three conditions for knowledge entails the other two.<sup>16</sup> Indubitability does not fit into the model in quite so neat a fashion, but of course it can be represented there. To say that one's beliefs in a certain area are immune from doubt is just to say that given any such belief, it is impossible for

<sup>12</sup> Recent criticism has shown that this analysis is not generally adequate without some modification. See, e.g., E. L. Gettier, "Is Justified True Belief Knowledge?" *Analysis*, vol. 23 (1962-63), pp. 121-123. However these difficulties do not attach to the sorts of cases with which we are concerned in this paper, and so we may take the above as a sufficient approximation for present purposes.

<sup>13</sup> It will be noted that in the previous discussion we were already presenting infallibility as an impossibility of error for one's *beliefs* or *judgments*.

<sup>14</sup> We use this cumbersome locution rather than the more natural "is infallible," so that our standard form will be usable for concepts such as indubitability that are not predicated of persons.

<sup>15</sup> Here too we might build up this formulation by first thinking of the fact that *S* (condition C.) entailing belief that *S* (condition A.), and then deriving the entailment of condition B. from that. For if a certain range of facts is such that it is impossible for such a fact to obtain without my believing that it does, it would seem clear that any such belief would be amply warranted. It would be amply warranted since it inevitably stems from the fact believed in.

<sup>16</sup> For a similar presentation, see D. M. Armstrong, *A Materialist Theory of the Mind* (London, 1968), p. 101.

anyone to have any grounds for doubting that the other two conditions for knowledge hold. Again, in the first instance, indubitability entails that there can be no grounds for doubting that one belief is true (condition C.); but if that is the case, then surely no one can have any grounds for doubting that one is justified in holding one's belief.

- (D3) *P* enjoys indubitability *vis-à-vis* a type of proposition,  $R =_{df}$ . For any proposition, *S*, of type *R*, it is logically impossible that *P* should believe that *S* and that anyone should have any grounds for doubting that *P* knows that *S*. (Condition A. for *P*'s knowing that *S* logically implies that there can be no grounds for doubting that conditions B. and C. hold.)

Incorrigibility can be given a similar formulation as follows:

- (D4) *P* enjoys incorrigibility *vis-à-vis* a type of proposition,  $R =_{df}$ . For any proposition, *S*, of type *R*, it is logically impossible that *P* should believe that *S* and that someone should show that *P* is mistaken in this belief. (Condition A. for *P*'s knowing that *S* logically implies that no one else can show that condition C. does not hold.)<sup>17, 18</sup>

Having now defined four different favorable epistemic positions in which a person may be *vis-à-vis* a given range of propositions, we can use these concepts to specify four ways in which a person may be said to have privileged access to his current mental states. To say that a person has *privileged* access to his current mental states is to say that his epistemic position *vis-à-vis* propositions ascribing current mental states to himself is favorable in a way no one else's position is. The simplest standard formula for a privileged access claim would be:

Each person enjoys ——— *vis-à-vis* propositions ascribing current mental states to himself, while no one else enjoys ——— *vis-à-vis* such propositions.

<sup>17</sup> It will be noted that all these definitions have been stated in terms of logical modalities. Later we shall explore the possibility of employing other modalities.

<sup>18</sup> We could, of course, make incorrigibility more parallel with the other modes by construing it to involve also the impossibility of anyone else's showing that *P* is not justified in believing that *S*. However, since this goes beyond what is either stated by our sources, or implied by what they say, I have avoided strengthening it in this way. It is clear that an impossibility of anyone else's showing that I am mistaken does not necessarily carry with it an impossibility of showing that my belief is unjustified. And this general possibility of dissociation might conceivably apply to beliefs about one's own mental states. It is conceivable, e.g., that one might show, through psychoanalysis, that I have a general tendency to deceive myself about my attitudes toward my daughter. This might well be taken to show that I am not justified in what I believe about those attitudes, even though no one is able to show (conclusively) that any particular belief I have about those attitudes is mistaken.

By successively filling in the blank with the four terms we have defined, we get four different privileged access theses.

However, it will be useful in our further discussions to have the different versions of a privileged access thesis spelled out more explicitly, with the content of the chosen mode of favorable epistemic position explicitly represented. We can give these more explicit formulations as follows:

- (T1) (Infallibility) Each person is so related to propositions ascribing current mental states to himself that it is logically impossible for him to believe that such a proposition is true without knowing it to be true; while no one else is so related to such propositions.
- (T2) (Omniscience) Each person is so related to propositions ascribing current mental states to himself that it is logically impossible for such a proposition to be true without his knowing that it is true; while no one else is so related to such propositions.
- (T3) (Indubitability) Each person is so related to propositions ascribing current mental states to himself that it is logically impossible both for him to believe that such a proposition is true and for anyone to have any grounds for doubting that he knows that proposition to be true; while no one else is so related to such propositions.
- (T4) (Incorrigibility) Each person is so related to propositions ascribing current mental states to himself that it is logically impossible both for him to believe that such a proposition is true and for someone else to show that that proposition is false; while no one else is so related to such propositions.

### III

As I pointed out earlier, it is not my aim in this work to determine in just what way, if any, one

does have privileged access to just what kinds of mental states. However it may help to motivate our consideration of other modes of privileged access if we briefly allude to some of the considerations that have led many thinkers to reject the modes so far considered. If we think of the range of mental states as including dispositional (belief, desire) as well as phenomenal (sensations, thoughts, feelings) states, there would seem to be strong reasons for denying that one enjoys infallibility, omniscience, indubitability, or incorrigibility *vis-à-vis* all the items within this range. The most dramatic reasons come from the sorts of cases highlighted by psycho-analysis, in which one hides certain of one's desires or beliefs from oneself, and in the process attributes to oneself desires or beliefs that one does not have. Thus consider the classic overprotective mother, who is preventing her daughter from going out in society in order to prevent her from developing into a feared rival. This mother stoutly and sincerely denies wanting to prevent her daughter's development and believing that her policy is likely to lead to any such result. Instead, she says, she is motivated solely by a desire to protect her daughter from harm. It certainly does seem at least possible that there are such cases in which the person both has desires and beliefs without knowing that he has them and attributes to himself desires and beliefs he does not have (at least not to the extent he supposes). Moreover, in such cases other people will have substantial grounds for doubting what the woman says about the desires and beliefs in question, and it even seems possible that others may sometimes be in a position to *show* (using realistic standards for this) that she is mistaken; so that not even indubitability or incorrigibility hold for beliefs and desires.

There is no doubt that proponents of these modes of privileged access are in a stronger position with respect to phenomenal states. I do not feel that this issue is definitely settled by a long way, but there are substantial negative arguments here.<sup>19</sup> For example, a general argument against infallibility is that knowledge of particular facts essentially involves the application of general concepts to those facts and hence is inherently liable to error. At the very least these negative arguments

provide a stimulus to consider whether there is not some weaker sense in which a person might be said to be in a necessarily superior epistemic position *vis-à-vis* his own mental states.

Another candidate that is well represented in the literature is "immediacy" or "directness." The notion that a person is privileged in having *immediate* knowledge of his own mental states is expressed incidentally in several of our original quotations. Thus Brentano says that the perception of mental phenomena is "immediately evident"; Ewing speaks of propositions based on introspection as "immediately known"; Hume speaks of our consciousness of the "actions and sensations of the mind" as that domain of experience where we are "most intimately conscious." Immediacy is closer to the center of the stage in the following quotations.

It has been suggested, namely, that any entity, which *can be directly known by one mind only* is a mental entity, and is "in the mind" of the person in question, and also, conversely, that all mental entities can be directly known only by a single mind.<sup>20</sup>

It is one such essential feature of what the word "mind" means that minds are private; that one's own mind is something with which one is directly acquainted—nothing more so—but that the mind of another is something which one is unable directly to inspect.<sup>21</sup>

The terms "immediate" and "direct" are susceptible of a variety of interpretations. Malcolm, in his essay, "Direct Perception,"<sup>22</sup> maintains that "impossibility of error" is the main feature of the standard philosophical conception of direct perception" (p. 89). He cites several eminent philosophers in support of this claim, including Berkeley, Moore, and Lewis. He then goes on to construct the following definition: "*A directly perceives x if and only if A's assertion that he perceives x could not be mistaken; . . .*" (*loc. cit.*). Of course if this is what we mean by directness, we have already discussed it under the heading of infallibility. We are therefore led to look for some other interpretation.

Moore, in typical fashion, tries to explain "direct knowledge" by pointing to a certain not further analyzable feature of our conscious experience. Immediately following the passage quoted above he writes:

<sup>19</sup> For some recent presentation of such arguments, see D. M. Armstrong, *op. cit.*, ch. 6, sect. 10; and B. Aune, *Knowledge, Mind, and Nature* (New York, 1967), ch. II, sect. 1.

<sup>20</sup> G. E. Moore, "The Subject Matter of Psychology," in Vesey, *op. cit.*, p. 241.

<sup>21</sup> C. I. Lewis, "Some Logical Considerations Concerning the Mental," in Vesey, *op. cit.*, p. 332.

<sup>22</sup> In *Knowledge and Certainty*, *op. cit.*

By "direct knowledge" is here meant the kind of relation we have to a colour, when we actually see it, or to a sound when we actually hear it.

But if we simply leave the matter there it is not very satisfactory. Presumably *something* can be said about the relation one has to a color when one actually sees it. And if it is not made explicit what the relation(s) in question is, we shall have no basis for resolving controversies over whether something or other is (or can be) directly known, whether, e.g., one directly knows that one has a certain belief, or whether it is conceivable that another person could directly know one's own thoughts. Let us try to find something more explicit.

Talk about immediate knowledge has traditionally been powerfully influenced by a spatial-causal model of mediacy. When people deny that perceptual knowledge of physical objects is immediate, it is often on the grounds that there is a spatial and causal gap between my knowledge (or rather the beliefs and/or sense-impressions involved) and the object of knowledge—the tree or whatever. There are spatial and causal intermediaries involved, and if these are not aligned properly, things can be thrown off. Similarly there are causal and spatial intermediaries between my desire or feeling and your belief that I have that desire or feeling. Your belief (in the most favorable case) is evoked by some perceptions of yours, which are in turn evoked by some behavior of mine, which is in turn evoked by my desire or feeling. But when it comes to my knowledge of my desires or feelings, no such intermediaries are involved, and here we do not have the same possibilities of distortion. I am "right up next to" my own mental states; I am "directly aware" of them; they give rise to my knowledge without going through a causal chain of any sort. Let us call immediacy so construed "causal immediacy."

The main reason for not using this sense of "immediacy" here is that we do not know how to determine either spatial or causal directness for knowledge of one's own mental states. We are not able to assign precise spatial locations to mental states.<sup>23</sup> Insofar as such location is possible, it is something rough, like "in the body," or "in the head," or maybe "in the brain." For the other cases that is enough for a judgment of mediacy; as long as my belief that there is a tree out there is somewhere in my head, it is clearly not spatially

contiguous to the tree; and as long as your desire is in your head and my belief that you have that desire is in my head, then they are not spatially contiguous (even if our heads are touching). But when both the belief and the object of the belief are mental states of the same person, we would need a more precise method of location to determine whether or not they are spatially contiguous. They are both "in the head", but just where in the head? Similar comments can be made concerning judgments of causal immediacy. With no more precise assumption than that the immediate causal antecedents of a belief of mine consist of processes in the brain, I can be sure that no belief of mine has its immediate causal antecedents in a tree. But if I am to determine whether my desire to go to Europe is an immediate causal antecedent of my belief that I have a desire to go to Europe, I need to have a more fine-grained view of the causal processes involved, and unfortunately we do not have any such view. We are in almost total ignorance of the causal processes, if any, involved in the origin of beliefs about our own mental states, and so we simply do not know what intermediaries there may be.

The upshot of this discussion is that although we can have sufficient reasons for terming many cases of knowledge "mediate" in the causal sense, we can have no assurance that any particular kind of knowledge is causally *immediate*, for when we come to the only plausible candidates for such immediacy, we do not know enough about the spatial and causal relations involved (if any) to have any basis for the denial of intermediaries. Thus our criterion is quite unworkable if we interpret it in terms of causal immediacy. It will be noted that we have argued for this without casting doubt on the intelligibility of the term "causal immediacy." Such a doubt could be raised, but that is another story.

There is a more distinctively epistemic sense of the mediate-immediate contrast, a sense that is suggested by such talk as the following. "You can know what I am thinking and feeling only *through* something (some signs, indications, criteria, or whatever); your knowledge of my thoughts and feelings is *based on* something else you know." But I, by contrast, know directly what I am thinking and feeling. I don't have to "derive" this knowledge from anything else. Let's say that in the sense of the contrast suggested by these remarks, mediate

<sup>23</sup> I am not maintaining, like some opponents of the identity theory, that such determinations are logically impossible. I am merely pointing out that at present we lack the resources for doing so.

knowledge is, while immediate knowledge is not, *based on* other knowledge.

However the term "based on" does not wear a unique interpretation on its face. It is often used in such a way that to say that my knowledge that *S* is based on my knowledge that *T* is to say that I arrived at the knowledge that *S* by inferring *S* from *T*. Thus philosophers have often used the presence or absence of inference as the crucial consideration in deciding whether a given piece of knowledge is to be called "direct" or "indirect."

I affirm, for example, that I hear a man's voice. This would pass, in common language, for a direct perception. All, however, which is really perception, is that I hear a sound. That the sound is a voice, and that voice the voice of a man, are not perceptions but inferences.<sup>24</sup>

However, it is clear that this contrast in terms of inference is not going to make the desired discriminations if we confine ourselves to conscious inference. The perception of speech does not ordinarily involve a conscious inference from the existence of a sound (under some acoustical description) to the existence of a human voice as its source. And more to the present point, it is clear that one's knowledge of the mental states of others is not always mediate if conscious inference is a necessary condition of mediacy. Quite often when I see that my companion feels dejected I am not aware of performing any inference from specifiable features of his speech, demeanor, and bearing to his dejection. And we certainly want to develop a concept of mediacy which is such that our ordinary knowledge of the mental states of others counts as mediate. Hence if we are going to make the desired discriminations in terms of the presence or absence of inference, we are going to have to rely heavily on the postulation of unconscious inference. I would not wish to subscribe to any general ban on such postulations, and it may be that we are justified in postulating unconscious inferences in just those cases where they are needed to discriminate between mediate and immediate knowledge along the present lines. However, in view of the obscurities surrounding the concept of unconscious inferential processes, and in view of present uncertainties concerning the conditions under which the postulations of such processes is justified, it would seem desirable to search for some other interpretation of "based on."

I would suggest that the tripartite analysis of knowledge introduced earlier provides us with the materials for such an interpretation. Using that schema we can distinguish between mediate and immediate knowledge in terms of what satisfies the second condition. If what justifies *P* in believing that *S* is some other knowledge that *P* possesses, then his knowledge is mediated by (based on) that other knowledge in a strictly epistemological sense. If, on the other hand, what satisfies condition (B.) is something other than *P*'s having some knowledge or other, we can say that his knowledge that *S* is *immediate*, not based on other knowledge. Let us call this kind of immediacy "epistemic immediacy." If I know that there was a fire last night at the corner of Huron and 5th because I read it in the *Ann Arbor News*, my knowledge is mediate; since what warrants me in believing that there is such a fire is my knowledge that such a fire was reported in the *Ann Arbor News*, plus my knowledge that it is a reliable source for local news. (The fact that the *Ann Arbor News* *did* carry this story and the fact that it *is* reliable in such matters will not justify *me* in believing that the fire took place, unless *I* know them to be the case.) Again if I know that my brother is dissatisfied with his job because he has complained to me about it, what warrants me in believing that he is dissatisfied is my knowledge that he has been complaining about it (and means what he says). On the other hand, it seems overwhelmingly plausible to suppose that what warrants me in believing that I feel disturbed, or am thinking about the mind-body problem, is not some other knowledge that I have. There is no bit of knowledge, or disjunction of bits of knowledge, such that if I do not have it (or some of them) my belief is not warranted. What would such bits of knowledge be? This is reflected in the oft-cited, but almost as often misunderstood,<sup>25</sup> fact that it "sounds odd," or even "nonsensical" to respond to a person who has just told us how he feels or what he is thinking, with "What reason do you have for saying that?" or "What is your evidence for that?" One does not know how to answer such a question; there is no answer to give.

This characterization of "immediate" is purely negative. It specifies what sort of thing does *not* satisfy condition B. where the knowledge is immediate, but it does not further limit the field of alternative possibilities. Clearly we can have

<sup>24</sup> J. S. Mill, *A System of Logic*, Bk. IV, ch. 1, sect. 2. Quoted by N. Malcolm, "Direct Perception," p. 88.

<sup>25</sup> It is misunderstood when it is taken to show that it makes no sense to speak of a person knowing that he feels disturbed, rather than taken to show what kind of knowledge this is.

different sorts of immediate knowledge claims depending on what is taken to satisfy condition B. If we consider the most explicit claims to immediate knowledge of one's own mental states in the literature, those of Ayer and Shoemaker, we shall see that in both cases condition C. is taken to imply condition B.

This gives us the clue also to what may be meant by saying that knowledge of this kind is direct. In other cases where knowledge is claimed, it is not sufficient that one be able to give a true report of what one claims to know: it is necessary also that the claim be authorized, and this is done by adducing some other statement which in some way supports the statement for which the claim is made. But in this case no such authority is needed; . . . Our knowledge of our thoughts and feelings accrues to us automatically in the sense that having them puts us in a position and gives us the authority to report them.<sup>26</sup>

. . . it is characteristic of a certain kind of statements, what I there called "first-person experience statements," that being entitled to assert such a statement does not consist in having established that the statement is true, i.e., in having good evidence that it is true or having observed that it is true, but consists simply in the statement's *being* true.<sup>27</sup>

Let us use the term "truth-sufficiency" for the sort of epistemic position described by these authors. We can put this notion into our standard format as follows:

- (D5) *P* enjoys truth-sufficiency *vis-à-vis* a type of proposition, *R* = *at*. For any true proposition, *S*, of type *R*, it is logically impossible that *P* should not be justified in believing that *S*. (Condition C. for *P*'s knowing that *S* logically implies condition B.)<sup>28</sup>

The privileged access thesis that makes use of this concept can be formulated as follows.

- (T5) (Truth-sufficiency) Each person is so related to propositions ascribing current mental states to himself that it is logically

impossible both for such a proposition to be true and for him not to be justified in believing it to be true; while no one else is so related to such propositions.

Knowledge involving truth-sufficiency is a sort of limiting case of direct knowledge; for here what is taken to justify the belief is something that is independently required for knowledge, viz., the truth of the belief. Thus nothing over and above the other two conditions for knowledge is required for the satisfaction of condition B., and so B. becomes, in a way, vacuous. We may call cases of knowledge in which nothing is required to satisfy B. over and above the other conditions for knowledge, "autonomous" knowledge.<sup>29</sup>

However, one can hold that a certain kind of knowledge is direct without considering it to be autonomous. Whenever condition B. is satisfied by something other than the possession of one or more pieces of other *knowledge* by the person in question, *and* this something goes beyond the other conditions for knowledge, we have knowledge that is direct but not autonomous. Thus a "direct realist," who denies that one's perceptual knowledge of physical objects is based on an epistemically prior knowledge of sense data, will think of perceptual knowledge as direct in the present sense of that term. However, he will certainly not think that nothing but the truth of *S* (a proposition describing a perceivable state of affairs) is required to justify a perceptual belief in *S*. The mere fact that it is true that there is now a fire in my living room fireplace does not justify me in believing this, and more specifically does not justify me in accepting it as a perceptual belief. I shall not be so justified if I am out of sensory range of the fire, if, e.g., there is a thick wall between me and the living room, or if my sense organs are not functioning properly. Thus justifiability will at least require the belief that *S* to have resulted from the normal operation of one's sense organs and central nervous system, as set into

<sup>26</sup> "Privacy," in *The Concept of a Person*, p. 64.

<sup>27</sup> S. Shoemaker, *op. cit.*, p. 216.

<sup>28</sup> This is the first time we have envisaged an implication of condition B. while the question of the satisfaction of condition A. is left undecided. It may seem that this is impossible, on the grounds that B. presupposes that A. is satisfied. How can I be justified in having a belief that I do not have? To make the three conditions logically independent, we shall have to interpret B. as: "P is in such a position that he will thereby be justified in believing that *S* if he has such a belief."

<sup>29</sup> Actually the concepts of infallibility and omniscience, as we have introduced them, satisfy our criteria for both directness and autonomy. If one is infallible or omniscient relative to a type of proposition, *R*, then when one knows that *S*, where *S* is an instance of that type, one's knowledge is both direct and autonomous. For what satisfied condition B. is, in the case of infallibility, condition A., and, in the case of omniscience, condition C. In considering autonomy as an alternative to infallibility and omniscience, we are restricting ourselves to cases in which one of the other conditions is sufficient for B., but not also sufficient for the third condition. We could, of course, build that further restriction into a definition of autonomy, but there will be no need to do so, since our list of modes of privileged access will not contain autonomy as such.



operation by stimuli from *S*. Perceptual knowledge so construed is direct but not autonomous.

In the light of these distinctions we can see that Ayer and Shoemaker have an inadequate conception of the alternatives to their version of direct knowledge. Let us recall that Ayer says:

In other cases where knowledge is claimed, it is not sufficient that one be able to give a true report of what one claims to know: it is necessary also that the claim be authorized, and this is done by adducing some other statement which in some way supports the statement for which the claim is made.

Ayer is contrasting (his version of) *autonomous* knowledge with mediate knowledge, ignoring the intermediate category of knowledge that is direct but not autonomous. As our reference to the direct realist view of perception shows, there may be additional "authorizations" required where these "authorizations" do not consist in the putative knower's having some other knowledge that can count as evidence for *S*. Shoemaker is a bit more inclusive; he gives as alternative modes of "entitlement," "having established that the statement is true, i.e., in having good evidence that it is true or having observed that it is true." The latter disjunct could presumably be construed so as to cover perceptual knowledge as viewed by the direct realist, though as stated the condition is uninformatively circular; to say that one has observed that it is true that *S* is just to say that one has perceptual knowledge that *S*. However, there are still many other possibilities for direct knowledge. For example, one might hold (with what justice I shall not inquire) that a belief about what makes for successful teaching is justified merely by the fact that one has engaged in a lot of teaching for a long time. More generally one may hold that long experience in an activity puts one in a position to make justified statements (of certain sorts) about that activity, regardless of whether one has any knowledge that could count as sufficient evidence for those statements.

Ayer and Shoemaker have not only overlooked the possibility of direct but non-autonomous knowledge; they have also failed to notice another possibility for autonomous knowledge, viz., taking A. instead of C. as a sufficient condition for B. To say that this is true of one's epistemic position *vis-à-vis* a certain range of propositions is to say that *any* belief in such a proposition is necessarily a justified one. We may use the term "self-warrant" for such a position.

(D6) *P* enjoys self-warrant *vis-à-vis* a type of proposition, *R* = *at*. For any proposition, *S*, of type *R*, it is logically impossible that that *P* should believe that *S* and not be justified in believing that *S*. (Condition A. for *P*'s knowing that *S* logically implies condition B.)

The corresponding privileged access thesis may be formulated as follows:

(T6) (Self-warrant) Each person is so related to propositions ascribing current mental states to himself that it is logically impossible both for him to believe that such a proposition is true and not be justified in holding this belief; while no one else is so related to such propositions.

It is clear that self-warrant and truth-sufficiency are weaker analogues of infallibility and omniscience, respectively. In the stronger modes, a given condition for knowledge is held to entail the other two, while in the weaker analogue that condition is held to entail *only* condition B. If one enjoys infallibility, then A. entails both B. and C., while with self-warrant, A. entails only B., leaving open the logical possibility of error. If one enjoys omniscience, C. entails both A. and B., while with truth-sufficiency C. entails only B., leaving open the logical possibility of ignorance.

Let us look more closely at the relations of self-warrant and truth-sufficiency. In a way they are equivalent. Both insure that conditions A. and C., which are required for knowledge in any event, are sufficient for any given piece of knowledge in the appropriate range. Whether I enjoy self-warrant or truth-sufficiency (or both) *vis-à-vis* my current thoughts and feelings, it will follow in either case that whenever I have a true belief to the effect that I am thinking or feeling *x* at the moment, I can correctly be said to *know* that I am thinking or feeling *x*. And neither privilege carries any guarantee that anything less will suffice for knowledge. However, they carry different implications as to what can be said short of a full knowledge claim. Enjoying self-warrant in this area guarantees that *any* belief of this sort is justified; it protects one against the possibility of unjustified belief formation. Whereas truth-sufficiency makes no such guarantee; it is compatible with the existence of some unjustified beliefs in the appropriate range. Does truth-sufficiency confer a contrasting partial advantage? Does it put the agent into some favorable position (short of knowledge) that he is not

put into by self-warrant? It may seem to. For it guarantees that for any thought or feeling possessed by  $P$  at  $t_1$ ,  $P$  is justified in believing that he currently has that thought or feeling. That is, with respect to whatever thought or feeling I have at a given time, the fact that I enjoy truth-sufficiency means that I possess the conditional guarantee that my belief that I currently have that thought or feeling will be justified *if* I have such a belief. But in fact this adds nothing to the guarantee given by self-warrant. For the latter involves the claim that *any* of  $P$ 's belief in the appropriate range, whether true or not, will (necessarily) be justified. Whereas truth-sufficiency guarantees this only for such beliefs as are true. The latter guarantee is a proper part of the former.<sup>30</sup>

Thus we may conclude that within the range of varieties of privileged access weaker than omniscience, infallibility, indubitability, and incorrigibility, self-warrant is the more interesting and important, since it provides everything in the way of cognitive superiority that is provided by truth-sufficiency, but not vice-versa. I would suggest that Ayer and Shoemaker missed the boat when they singled out truth-sufficiency for consideration.

The greater interest of self-warrant is also shown by its greater utility as a principle of cognitive evaluation. We are now taking the standpoint of another person evaluating  $P$ 's knowledge claims, rather than the standpoint of  $P$  and his cognitive capacities. The basic point is that the criteria of justification provided by self-warrant are more accessible than those in terms of which the truth-sufficiency principle is stated. It is generally much easier to determine whether  $P$  believes that he has a certain thought, than it is to determine whether in fact he does have that thought. At least that is the case, insofar as a determination of the latter is

a task that goes beyond the determination of the former. And of course where we are employing truth-sufficiency *instead of* self-warrant as a principle for the evaluation of knowledge claims, we must be taking the verification of condition C. to be distinguishable from the verification of condition A.; otherwise the use of the truth-sufficiency principle could not be distinguished from the use of the self-warrant principle.

#### IV

I believe that (T6) is the most defensible of the privileged access principles we have considered. It escapes the objections urged against claims of infallibility, omniscience, indubitability, and incorrigibility. It allows for cases in which a person is mistaken about his current mental states, (and of course it puts no limit at all on the extent to which a person may be ignorant of his current mental states), and it even allows for cases in which someone else can show that one is mistaken. And at the same time it specifies a very definite respect in which a person is in a superior epistemic position *vis-à-vis* his own mental states. To be sure, it is not immune from criticism. A thorough examination of such criticism is outside the scope of this paper, but there is one plausible criticism a consideration of which will afford a convenient entrée to still further varieties of privileged access.

The criticism in question is an attack on the logical entailment (and logical impossibility) claim that is imbedded in (T6). It maintains that what the principle holds to be logically impossible, viz., that a first-person-current-mental-state-belief (FPCMSB) should be unwarranted, is in fact consistently conceivable.<sup>31</sup> More specifically, it claims that it is just a matter of fact that people are

<sup>30</sup> In fact, if we should interpret truth-sufficiency as involving the claim that C. is a necessary as well as a sufficient condition for the justification of the belief (for B.), then not only does truth-sufficiency not confer any additional cognitive advantage over self-warrant; it puts one in a less advantageous position. For, contrary to self-warrant, it entails that one cannot be justified in a belief about one's own mental states unless that belief is true. That is, it entails a reduction in the range of cases (intensionally even if not extensionally) in which one's beliefs are justified.

My formulation of truth-sufficiency did not represent C. as a necessary condition for B., though Shoemaker could be interpreted in this way. Of course for Shoemaker it is hardly a live issue, since he also commits himself to infallibility without explicitly distinguishing it from truth-sufficiency. If one accepts the infallibility principle, then the question whether C. is necessary as well as sufficient for B. becomes otiose. For since A. entails C. there is not even a logical possibility of a case in which one would justifiably believe that  $S$  (A. and B.) without  $S$ 's being true (C.). However in the case of someone like Ayer, who rejects infallibility, it would seem unjustifiable to make C. a necessary condition for B. For this would be to put a person in a worse cognitive position, in a way, with respect to his own mental states than he is in other fields of knowledge. For when it comes to knowledge of the physical world and historical events it is possible to be justified in believing that  $S$ , even though it is not true that  $S$ .

<sup>31</sup> There are also arguments to the effect that it is logically possible for other persons to enjoy self-warrant (and/or other modes of favorable epistemic position) *vis-à-vis* one's own mental states; but I will not have time to go into that side of the criticism.

highly reliable in the reports they give us about their feelings, thoughts, and beliefs, i.e., that things generally turn out as they could be expected to on the hypothesis that those reports are correct. As things are, if a person tells us that he is feeling depressed or that he is thinking about his income tax, then subsequent events tend to bear this out, insofar as we can form any definite rational expectations, given the complexity of the connections, the undeveloped state of our knowledge of general connections in this area, and the varied possibilities for dissimulation. However, it is quite conceivable that the world should be such that a person's reports of his feelings, thoughts, and beliefs would be no better guide to the future than, say, his reports of his immediate physical environment, which are still highly reliable, but by no means so overbearingly so as to be rightly accepted as self-warranted. If the world were like this, it would depend on further factors whether a given FPCMSB were warranted, just as is now the case with perceptual beliefs. Such factors might include how alert the person is at the moment, how good a judge he has proved himself to be in such matters, and so on. Thus in this logically possible world FPCMSB's would be sometimes unwarranted, viz., in those cases in which the requisite additional factors were not present. But then even in our world it cannot be *logically* impossible for a FPCMSB to be unwarranted.<sup>32</sup>

There are various ways in which a defender of self-warrant may try to meet this criticism. First he may roundly deny that what the criticism maintains to be logically impossible is indeed so. This denial may take varying forms, but I would suppose that the most plausible is one based on the claim that as we now use mental state terms like "feel ———," "think about ———," etc., it is "part of their meaning" that FPCMSB's are self-warranted. It is impossible, in *our language*, to make sense of the supposition that such a belief should not be warranted. What the critic is doing, in effect, is envisaging a situation in which the meanings of mental state terms would have changed in this respect. But the fact that such a change in meaning is possible leaves untouched the point that, given the present meaning of mental-state terms, having

a FPCMSB logically guarantees a warrant for the belief.<sup>33</sup>

At the other end of the spectrum would be a capitulation; one might concede, in the face of the criticism, that the epistemic superiority enjoyed by FPCMSB's is only an "empirical" one; it is just a matter of fact that one's own reports of one's current mental states are more reliable in general than the estimates of those states formed by other people.

For present purposes we are interested in replies that lie between these extremes, replies that develop conceptions of self-warrant that lie between a *logical* impossibility of unwarranted belief and a mere *de facto* superior reliability. For this purpose let us imagine that the self-warrant theorist agrees with the critic that the situations adduced by the latter can be consistently described, using terms with their current meanings. Thus he has to give up the claim that FPCMSB's logically entail their own warrant. Nevertheless he still feels inclined to assert a stronger kind of superiority of such beliefs over their third-person counterparts than merely a greater frequency of accuracy.

An obvious move at this point is to consider the possibility of defining self-warrant in terms of modalities other than logical. Two kinds that are familiar from other contexts are nomological modalities and normative modalities. The former is illustrated by such sentences as:

Water can't run uphill.

If the cream's been around that long it has to be sour.

An airplane *could* go 1200 miles an hour.

Here the modalities are based on laws of nature rather than on logical principles. There is nothing logically impossible about water running uphill, but to do so would be contrary to physical laws. The sourness of the cream in question follows from biochemical laws plus antecedent conditions. And so on.

Normative modalities are employed in sentences like the following:

Bringing happiness to another person can't be a sin.

To get a Ph.D. you have to write a thesis.

You can't win without scoring runs.

<sup>32</sup> Such a criticism could, of course, also be brought against the other varieties of privileged access we have distinguished since they all were stated in terms of logical entailment and logical impossibility. However, I have chosen to state the criticism in opposition to self-warrant, since I take this to be the strongest form of privileged access that does not fall victim to other objections.

<sup>33</sup> Just as the criticism in question may be made of stronger forms of a privileged access principle, so this sort of reply may be made in defense of those stronger forms. Malcolm and Shoemaker seem to think that infallibility is guaranteed by the meaning we now attach to mental state terms.

Here the necessities and impossibilities are based on normative principles of one sort or another; in these examples they are moral standards, institutional regulations, and rules of games, respectively.

Can we conceive the impossibility of a FPCMSB's being unwarranted in one of these other ways, rather than as logical? It would seem that nomological impossibility is inapplicable here. If it were applicable, the laws involved would presumably be psychological laws, more specifically laws governing the formation of beliefs about one's own mental states. If these laws either asserted or implied a universal connection between a belief's being about one's own mental states and that belief's being warranted, then it would be nomologically impossible for such a belief to be unwarranted. The trouble with this is that warrantedness, being a normative concept, is not of the right sort to figure in a scientific law. To say that a belief is warranted is to say that it comes up to the proper epistemic standards, and to determine what the proper standards are for one or another kind of belief is not within the province of an empirical science, anymore than is any other normative question.<sup>34</sup> Hence it could not be nomologically impossible for a belief to be unwarranted.

But for just the same reasons normative modalities are quite appropriate here. If there is a justifiable epistemic norm to the effect that any FPCMSB is *ipso facto* warranted just by virtue of its being the belief it is, then it would thereby be normatively impossible for such a belief to be unwarranted. Thus an alternative formulation of the self-warrant principle would be:

(T6A) Each person is so related to propositions ascribing current mental states to himself that it is normatively impossible both for him to believe that such a proposition is true and not be justified in holding this belief; while no one else is so related to such propositions.

(T6A) presupposes that a defensible epistemic standard would effect a direct connection between being a FPCMSB and being warranted; hence it can be stated in terms of normative modalities alone. A more complicated case can also be envisaged, one in which both nomological and

normative modalities are involved. Suppose that there is no such defensible epistemic standard as the one just mentioned; suppose instead that there is a standard to the effect that any belief possessing a certain property, *F*, is warranted, *and* that it is nomologically necessary that all FPCMSB's have the property *F*. In that case it would be nomologically-normatively impossible for FPCMSB's to be unwarranted; the impossibility would derive from a combination of epistemic norms and psychological laws. This gives rise to another version of the self-warrant principle:

(T6B) Each person is so related to propositions ascribing current mental states to himself that it is nomologically impossible for him to believe such a proposition without his belief having property *F*, and it is normatively impossible for any belief having the property *F* to be unjustified; while no one else is so related to such propositions.

Both of these alternative versions of a self-warrant principle are immune to the criticisms that launched us on the search for alternative modalities. Since neither version asserts any logical necessity for the warrant of FPCMSB's, they allow for the possibility of consistently describing, in our present language, a situation in which (some) FPCMSB's would not be warranted. According to (T6A) such a situation would involve some difference in epistemic standards, and according to (T6B) it would involve either that or a difference in psychological laws. But there is nothing in these principles to suggest that such differences cannot be described by the use of current mentalistic language.

This discussion suggests the possibility of a considerable proliferation of privileged access theses, through varying the modalities in terms of which each of our six modes of favorable epistemic position is stated. However we should not suppose that every such mode can be construed in terms of every modality. We have already seen that self-warrant is not amenable to statement in terms of a purely nomological modality. The question of just which modalities are combinable with each mode of favorable epistemic position is a complicated one, and we shall have time only for a few sketchy and

<sup>34</sup> This conclusion is of course controversial, and it may be contested on the grounds that it presupposes an unwarranted distinction between the scientific (factual) and normative. However, it is not essential to this paper to take a definite stand on this issue. If normative matters are deemed to be within the province of science, it just means that nomological modalities are more widely applicable than I am supposing here, and the varieties of privileged access are even more numerous than I am representing them to be.

dogmatic remarks. It would seem that indubitability, incorrigibility, and truth-sufficiency cannot be stated in terms of nomological modalities alone, for the same reasons we gave in the case of self-warrant. With truth-sufficiency, which has the same structure, exactly the same argument applies. Our concept of indubitability being a normative one, it seems clear that an appeal to epistemic standards is essentially involved in the claim that there can be no *grounds* for any doubt. And since incorrigibility involves the impossibility of someone else's showing the person to be mistaken (not just the impossibility of someone else's being correct while one is mistaken), again it would seem that standards defining what counts as a demonstration in this area would be involved. By similar arguments I think it could be shown that these three modes are all construable, like self-warrant, in terms of both pure normative modalities and a combination of nomological and normative modalities.

The situation is quite different with respect to infallibility and omniscience. Here what is implied most basically is the correctness of a belief, and the possession of a belief, respectively. Whether or not a certain (factual) belief is correct, and whether or not a given person has a given belief, is sheerly a matter of fact and not within the jurisdiction of norms or standards. It would not make sense to adopt a standard to the effect that all beliefs of a certain category are correct; that would be as if a city council were to adopt an ordinance according to which all public housing in the city is free of rats. If a certain belief is in fact mistaken, we cannot alter that fact by legislating it away. On the other hand, it might conceivably be *nomologically* necessary that beliefs of a certain category all be correct. The mechanisms of belief formation might be such as to guarantee this. However, since both infallibility and omniscience, as we have formulated them, involve the implication of the warrant condition, B., as well, it would seem, for the reasons given in connection with self-warrant, that they cannot involve nomological modalities alone.<sup>35</sup> Thus a mixed nomological-normative modality would seem to be the only alternative to the logical modalities (at least among the alternatives we are

considering). An infallibility principle could be stated in those terms as follows:

- (T1A) Each person is so related to propositions ascribing current mental states to himself that it is nomologically impossible for him to believe such a proposition without that proposition's being true, and it is normatively impossible for a belief that satisfies this condition to be unjustified; while no one else is so related to such propositions.

In addition, there is the possibility of non-modal, *de facto* universal, versions of each of our six basic types. The following is a rough indication of how privileged access principles will look as so construed.

Infallibility—FPCMSB's are, in fact, never mistaken.

Omniscience—A person is, in fact, never ignorant of one of his mental states.

Indubitability—No one, in fact, ever has grounds for doubting a FPCMSB.

Incorrigibility—No one else ever, in fact, succeeds in showing that a FPCMSB is mistaken.

Self-warrant—FPCMSB's are, as a matter of fact, always warranted.

Truth-sufficiency—True FPCMSB's are, as a matter of fact, always warranted.

It is dubious that these formulations in terms of *de facto* universals are of much use. For it is doubtful that anyone could have solid grounds for supposing that, e.g., FPCMSB's are *always* correct, unless his claim were based on some sort of modal consideration to the effect that something in the nature of the case makes it impossible for FPCMSB's to be mistaken; and in that case he would be in a position to formulate the infallibility thesis in modal terms. We can derive a more usable non-modal formulation by weakening the universals to a "for the most part" status, thus replacing "never" with "rarely," and "always" with "usually." So construed, infallibility and self-warrant, e.g., would become:

Infallibility—FPCMSB's are, in fact, rarely mistaken.<sup>36</sup>

<sup>35</sup> Of course we could excise this implication from our construal of infallibility and omniscience, in which case infallibility would be defined simply as the impossibility of mistake and omniscience simply as the inevitability of belief-formation. (It would still be the case that the beliefs that satisfy these conditions would be warranted and so count as knowledge; it is just that this would not be made explicit in the definition.) As so conceived these modes would be amenable to a formulation in terms of nomological modalities alone.

<sup>36</sup> Absolutistic terms like "infallibility" and "omniscience" are not aptly used for formulations that are weakened to this extent, but that is a merely verbal point.

Self-warrant—FPCMSB's are, as a matter of fact, usually warranted.

It does seem *conceivable* that one should acquire solid grounds for, e.g., a "for the most part" version of an infallibility principle without thereby putting himself in a position to assert a corresponding modal principle. One might argue, e.g., that we are rarely, if ever, in a position to show that a FPCMSB is incorrect, and that to the extent that we have an independent check on their accuracy, they almost always turn out to be correct. Therefore we have every reason to suppose that they are, in fact, usually correct. This line of argument is not based on any fundamental considerations concerning the concepts, natural laws, or epistemic norms involved, and therefore it does not support any claim as to what is necessary or impossible.

Another way of deviating from unrestricted universality involves a restriction to "normal" conditions. Thus we might construe self-warrant, e.g., as follows:

Self-warrant—FPCMSB's are normally warranted (are always warranted in normal conditions).

Certain lines of argument support a "normal conditions" version instead of, or in addition to, a "by and large" version. For example, one may think that he can give an (open-ended) list of "abnormal" factors that are usually absent but which when present would prevent a FPCMSB from having the warrant it usually has. These might include such things as extreme preoccupation with other matters, extreme emotional upset, and derangement of the critical faculties.

Our list of varieties of privileged access has now swollen to a staggering 34: 16 modal principles, 6 *de facto* unrestrictedly universal principles, 6 "by and large" principles, and 6 "normal conditions" principles. No doubt with sufficient ingenuity the list could be further expanded, but perhaps the results already attained will suffice to bring out the main dimensions of variation.

## V

Attacks on privileged access invariably fail to take account of the full range of possibilities. Their arguments are directed against only some of the possible versions of the position they are attacking;

hence at best they fall short of showing that no privileged access principle is acceptable. To illustrate this point I shall take two of the best and most prominent recent attacks on privileged access, those of Bruce Aune in his book *Knowledge, Mind, and Nature*,<sup>37</sup> and those of D. M. Armstrong in his book *A Materialist Theory of the Mind*.<sup>38</sup>

Aune's arguments occur in chap. II, "Does Knowledge Have An Indubitable Foundation?" and they form part of his attack on the general idea that our knowledge rests on a foundation that is made up of beliefs, each of which is wholly non-inferential and completely infallible. Although at the beginning of the discussion (pp. 32-33) he wobbles a bit between talking of the "reliability of our beliefs concerning immediate experience" (p. 32) and talking about "the alleged infallible character of immediate awareness," (p. 33) it is clear that his actual arguments are directed against an infallibility thesis, and indeed a logical infallibility thesis. The arguments are designed to show that "identifications of even feelings and mental images are not logically incapable of error" (p. 33), and to demonstrate "the possibility of being mistaken about the character of one's momentary experience" (p. 34). But having presented what he claims to be possible cases of mistake about such matters, he then supposes himself to have shown that statements about one's immediate experience are not "intrinsically acceptable," in the sense that their "truth is acceptable independently of any inference" (p. 41). And he seems to suppose that the only alternative to an infallibility claim is the view that there are strong empirical reasons for accepting one's statements about his immediate experience.

The point seems to be securely established that judgments of phenomenal identification are not, in fact, infallible. We may come to have enormous confidence that, after a protracted period of training, a man's opinions about the character of his own experiences are never really wrong. But our confidence here is based on empirical considerations. There is no longer any reason to think that such opinions *cannot* be erroneous; rather we have fairly good, though not infallible, reasons to think that they are normally reliable. (P. 37).

In supposing that having disposed of logical infallibility, he has thereby disposed of "intrinsic acceptability" and left a clear field for an acceptability based on empirical evidence, it is clear that

<sup>37</sup> New York, 1967.

<sup>38</sup> London, 1968.

Aune has overlooked most of the modes of epistemic superiority we have distinguished, in particular self-warrant and truth-sufficiency. For to say that one enjoys self-warrant *vis-à-vis* his own immediate experience is to say that any of the person's beliefs about that experience is justified (acceptable) just by virtue of the fact that it is held, whether or not one possesses anything that has the status of evidence for that belief; and to say that one enjoys truth-sufficiency *vis-à-vis* one's immediate experience is to say the same thing for one's true beliefs about that experience. These are two forms of "intrinsic acceptability" without infallibility, and Aune will have to mount arguments against these before he can lay claim to having disposed of "intrinsic acceptability." At the most basic level what is overlooked here is the distinction between a belief's guaranteeing its own truth, and a belief's guaranteeing its own justification (acceptability), i.e., the difference between the justification (B.) and the truth (C.) conditions for knowledge.

Moreover Aune has overlooked the possibility of working, within any of those modes of epistemic superiority, with modalities other than the logical modalities. In the quotation just cited, he presents a "normal conditions" variety of infallibility as the only alternative to a logical infallibility thesis. But even within the bounds of infallibility there is also the view that it is *nomologically* impossible that FPCMSB's should be mistaken. And as for self-warrant, even if we should reject the claim that it is logically impossible for FPCMSB's to be unwarranted, there is still the view that their being unwarranted is a normative impossibility, a view that is also distinct from the position that FPCMSB's are "normally reliable."

Armstrong's discussion (in chap. 6, sect. x of the above mentioned book, entitled, "The alleged

indubitability of consciousness") draws more distinctions than Aune. He distinguishes logical infallibility, logical omniscience, and logical incorrigibility (pp. 101-102), in much the same way as I, though he uses for these concepts the terms "indubitability" (alternatively "incorrigibility"), "self-intimation," and "logically privileged access," respectively. He then presents arguments against the claims that FPCMSB's enjoy any of these kinds of privilege. Furthermore he explicitly recognizes nomological infallibility.<sup>39</sup> In chap. 9, sect. ii, entitled "The nature of non-inferential knowledge" he defines non-inferential knowledge as follows: "*A* knows *p* non-inferentially, if, and only if, *A* has no good reasons for *p* but:

- (1) *A* believes *p*;
- (2) *p* is true;
- (3) *A*'s belief-that-*p* is empirically sufficient for the truth of *p*." (P. 189.)

He goes on to make it explicit that the "empirical sufficiency" involved here is based on some law of nature (p. 190); hence whenever one has non-inferential knowledge of *p* it is nomologically impossible that his belief should be mistaken. Thus he does not, like Aune, regard some kind of *de facto* reliability as the only alternative to a logically necessary epistemic privilege. However, like Aune, he fails to note both the possibility of self-warrant (and truth-sufficiency) as distinct modes of epistemic privilege, and the possibility of formulating at least some of the modes in terms of normative modalities. Thus even if he has effectively disposed of logical infallibility, omniscience, and incorrigibility, he cannot yet conclude that nomological infallibility is the only plausible version of a modal privileged access thesis.

*The University of Michigan*

*Received July 27, 1970*

<sup>39</sup> Where this is construed simply as matter of condition *A*'s nomologically implying *C* without the further stipulation of an implication of *B*.