On Rules and Private Language 37

SAUL KRIPKE

. . . A common view of the 'private language argument' in *Philosophical Investigations* assumes that it begins with section 243, and that it continues in the sections immediately following.[1] This view takes the argument to deal primarily with a problem about 'sensation language'. Further discussion of the argument in this tradition, both in support and in criticism, emphasizes such questions as whether the argument invokes a form of the verification principle, whether the form in question is justified, whether it is applied correctly to sensation language, whether the argument rests on an exaggerated scepticism about memory, and so on. Some crucial passages in the discussion following §243—for example, such celebrated sections as §258 and §265—have been notoriously obscure to commentators, and it has been thought that their proper interpretation would provide the key to the 'private language argument'.

In my view, the real 'private language argument' is to be found in the sections *preceding* §243. Indeed, in §202 *the conclusion is already stated explicitly:* "Hence it is not possible to obey a rule 'privately': otherwise thinking one was obeying a rule would be the same thing as obeying it." I do not think that Wittgenstein here thought of himself as *anticipating* an argument he was to give in greater detail later. On the contrary, the crucial considerations are all contained in the discussion leading up to the conclusion stated in §202. The sections following §243 are meant to be read in the light of the preceding discussion; difficult as they are in any case, they are much less likely to be understood if they are read in isolation. The 'private language argument' as applied to *sensations* is only a special case of much more general considerations about language previously argued; sensations have a crucial role as an (apparently) convincing *counterexample* to the general considerations previously stated. Wittgenstein therefore goes over the ground again in this special case, marshalling new specific considerations appropriate to it. It should be borne in mind that *Philosophical Investigations* is not a systematic philosophical work where conclusions, once definitely established, need not be reargued. Rather the *Investigations* is written as a perpetual dialectic, where persisting worries, expressed by the voice of the imaginary interlocutor, are never definitively silenced. Since the work is not presented in the form of a deductive argument with definitive theses as conclusions, the same ground is covered repeatedly, from the point of view of various special cases and from different angles, with the hope that the entire process will help the reader see the problems rightly.

The basic structure of Wittgenstein's approach can be presented briefly as follows: A certain problem, or in Humean terminology, a 'sceptical paradox', is presented concerning the notion of a rule. Following this, what Hume would have called a 'sceptical solution'

From *On Rules and Private Language* (Cambridge: Harvard University Press, 1982).

to the problem is presented. There are two areas in which the force, both of the paradox and of its solution, are most likely to be ignored, and with respect to which Wittgenstein's basic approach is most likely to seem incredible. One such area is the notion of a mathematical rule, such as the rule for addition. The other is our talk of our own inner experience, of sensations and other inner states. In treating both these cases, we should bear in mind the basic considerations about rules and language. Although Wittgenstein has already discussed these basic considerations in considerable generality, the structure of Wittgenstein's work is such that the special cases of mathematics and psychology are not simply discussed by citing a general 'result' already established, but by going over these special cases in detail, in the light of the previous treatment of the general case. By such a discussion, it is hoped that both mathematics and the mind can be seen rightly: since the temptations to see them wrongly arise from the neglect of the same basic considerations about rules and language, the problems which arise can be expected to be analogous in the two cases. In my opinion, Wittgenstein did not view his dual interests in the philosophy of mind and the philosophy of mathematics as interests in two separate, at best loosely related, subjects, as someone might be interested both in music and in economics. Wittgenstein thinks of the two subjects as involving the same basic considerations. For this reason, he calls his investigation of the foundations of mathematics "analogous to our investigation of psychology" (p. 232). It is no accident that essentially the same basic material on rules is included in both *Philosophical Investigations* and in *Remarks on the Foundations of Mathematics,*[2] both times as the basis of the discussions of the philosophies of mind and of mathematics, respectively, which follow.

In the following, I am largely trying to present Wittgenstein's argument, or, more accurately, that set of problems and arguments which I personally have gotten out of reading Wittgenstein. With few exceptions, I am *not* trying to present views of my own; neither am I trying to endorse or to criticize Wittgenstein's approach. In some cases, I have found a precise statement of the problems and conclusions to be elusive. Although one has a strong sense that there is a problem, a rigorous statement of it is difficult. I am inclined to think that Wittgenstein's later philosophical style, and the difficulty he found (see his preface) in welding his thought into a conventional work presented with organized arguments and conclusions, is not simply a stylistic and literary preference, coupled with a *penchant* for a certain degree of obscurity,[3] but stems in part from the nature of his subject.

I suspect—for reasons that will become clearer later—that to attempt to present Wittgenstein's argument precisely is to some extent to falsify it. Probably many of my formulations and recastings of the argument are done in a way Wittgenstein would not himself approve. So the present paper should be thought of as expounding neither 'Wittgenstein's' argument nor 'Kripke's': rather Wittgenstein's argument as it struck Kripke, as it presented a problem for him.

As I have said, I think the basic 'private language argument' *precedes* section 243, though the sections following 243 are no doubt of fundamental importance as well. I propose to discuss the problem of 'private language' initially without mentioning these latter sections *at all.* Since these sections are often thought to *be* the 'private language argument', to some such a procedure may seem to be a presentation of Hamlet without the prince. Even if this is so, there are many other interesting characters in the play.

. . . In §201 Wittgenstein says, "this was our paradox: no course of action could be determined by a rule, because every course of action can be made to accord with the rule." In this section of the present essay, in my own way I will attempt to develop the "paradox" in question. The "paradox" is perhaps the central problem of *Philosophical Investigations.* Even someone who disputes the conclusions regarding 'private language', and the philosophies of mind, mathematics, and logic,

that Wittgenstein draws from his problem, might well regard the problem itself as an important contribution to philosophy. It may be regarded as a new form of philosophical scepticism.

Following Wittgenstein, I will develop the problem initially with respect to a mathematical example, though the relevant sceptical problem applies to all meaningful uses of language. I, like almost all English speakers, use the word "plus" and the symbol '+' to denote a well-known mathematical function, addition. The function is defined for all pairs of positive integers. By means of my external symbolic representation and my internal mental representation, I 'grasp' the rule for addition. One point is crucial to my 'grasp' of this rule. Although I myself have computed only finitely many sums in the past, the rule determines my answer for indefinitely many new sums that I have never previously considered. This is the whole point of the notion that in learning to add I grasp a rule: my past intentions regarding addition determine a unique answer for indefinitely many new cases in the future.

Let me suppose, for example, that '68 + 57' is a computation that I have never performed before. Since I have performed—even silently to myself, let alone in my publicly observable behavior—only finitely many computations in the past, such an example surely exists. In fact, the same finitude guarantees that there is an example exceeding, in both its arguments, all previous computations. I shall assume in what follows that '68 + 57' serves for this purpose as well.

I perform the computation, obtaining, of course, the answer '125'. I am confident, perhaps after checking my work, that '125' is the correct answer. It is correct both in the arithmetical sense that 125 is the sum of 68 and 57, and in the metalinguistic sense that "plus," as I intended to use that word in the past, denoted a function which, when applied to the numbers I called "68" and "57," yields the value 125.

Now suppose I encounter a bizarre sceptic. This sceptic questions my certainty about my answer, in what I just called the 'metalinguistic' sense. Perhaps, he suggests, as I used the term "plus" in the past, the answer I intended for '68 + 57' should have been '5'! Of course the sceptic's suggestion is obviously insane. My initial response to such a suggestion might be that the challenger should go back to school and learn to add. Let the challenger, however, continue. After all, he says, if I am now so confident that, as I used the symbol '+', my intention was that '68 + 57' should turn out to denote 125, this cannot be because I explicitly gave myself instructions that 125 is the result of performing the addition in this particular instance. By hypothesis, I did no such thing. But of course the idea is that, in this new instance, I should apply the very same function or rule that I applied so many times in the past. But who is to say what function this was? In the past I gave myself only a finite number of examples instantiating this function. All, we have supposed, involved numbers smaller than 57. So perhaps in the past I used "plus" and '+' to denote a function which I will call 'quus' and symbolize by '⊕'. It is defined by:

$$x \oplus y = x + y \text{ if } x, y < 57$$
$$= 5 \text{ otherwise.}$$

Who is to say that this is not the function I previously meant by '+'?

The sceptic claims (or feigns to claim) that I am now misinterpreting my own previous usage. By "plus," he says, I *always meant* quus;[4] now, under the influence of some insane frenzy, or a bout of LSD, I have come to misinterpret my own previous usage.

Ridiculous and fantastic though it is, the sceptic's hypothesis is not logically impossible. To see this, assume the common sense hypothesis that by '+' I *did* mean addition. Then it would be *possible,* though surprising, that under the influence of a momentary 'high', I should misinterpret all my past uses of the plus sign as symbolizing the quus function, and proceed, in conflict with my previous linguistic intentions, to compute 68 plus 57 as 5. (I would have made a mistake, not in mathematics, but in the supposition that I had accorded with my previous linguistic intentions.) The sceptic is proposing that I have

made a mistake precisely of this kind, but with a plus and quus reversed.

Now if the sceptic proposes his hypothesis sincerely, he is crazy; such a bizarre hypothesis as the proposal that I always meant quus is absolutely wild. Wild it indubitably is, no doubt it is false; but if it is false, there must be some fact about my past usage that can be cited to refute it. For although the hypothesis is wild, it does not seem to be a priori impossible.

Of course this bizarre hypothesis, and the references to LSD, or to an insane frenzy, are in a sense merely a dramatic device. The basic point is this. Ordinarily, I suppose that, in computing '68 + 57' as I do, I do not simply make an unjustified leap in the dark. I follow directions I previously gave myself that uniquely determine that in this new instance I should say '125'. What are these directions? By hypothesis, I never explicitly told myself that I should say '125' in this very instance. Nor can I say that I should simply 'do the same thing I always did,' if this means 'compute according to the rule exhibited by my previous examples.' That rule could just as well have been the rule for quaddition (the quus function) as for addition. The idea that in fact quaddition *is* what I meant, that in a sudden frenzy I have changed my previous usage, dramatizes the problem.

In the discussion below the challenge posed by the sceptic takes two forms. First, he questions whether there is any *fact* that I meant plus, not quus, that will answer his sceptical challenge. Second, he questions whether I have any reason to be so confident that now I should answer '125' rather than '5'. The two forms of the challenge are related. I am confident that I should answer '125' because I am confident that this answer also accords with what I *meant*. Neither the accuracy of my computation nor of my memory is under dispute. So it ought to be agreed that *if* I meant plus, then unless I wish to change my usage, I am justified in answering (indeed compelled to answer) '125', not '5'. An answer to the sceptic must satisfy two conditions. First, it must give an account of what fact it is (about my mental state) that constitutes my meaning plus, not quus. But further, there is a condition that any putative candidate for such a fact must satisfy. It must, in some sense, show how I am justified in giving the answer '125' to '68 + 57'. The 'directions' mentioned in the previous paragraph, that determine what I should do in each instance, must somehow be 'contained' in any candidate for the fact as to what I meant. Otherwise, the sceptic has not been answered when he holds that my present response is arbitrary. Exactly how this condition operates will become much clearer below, after we discuss Wittgenstein's paradox on an intuitive level, when we consider various philosophical theories as to what the fact that I meant plus might consist in. There will be many specific objections to these theories. But all fail to give a candidate for a fact as to what I meant that would show that only '125', not '5', is the answer I 'ought' to give.

The ground rules of our formulation of the problem should be made clear. For the sceptic to converse with me at all, we must have a common language. So I am supposing that the sceptic, provisionally, is not questioning my *present* use of the word "plus"; he agrees that, according to my *present* usage, '68 plus 57' denotes 125. Not only does he agree with me on this, he conducts the entire debate with me in my language as I *presently* use it. He merely questions whether my present usage agrees with my past usage, whether I am *presently* conforming to my *previous* linguistic intentions. The problem is not "How do I know that 68 plus 57 is 125?", which should be answered by giving an arithmetical computation, but rather "How do I know that '68 plus 57', as I *meant* 'plus' in the *past*, should denote 125?" If the word "plus" as I used it in the past, denoted the quus function, not the plus function ('quaddition' rather than addition), then my *past* intention was such that, asked for the value of '68 plus 57', I should have replied '5'.

I put the problem in this way so as to avoid confusing questions about whether the discussion is taking place 'both inside and outside language' in some illegitimate sense.[5] If we are querying the meaning of the word "plus," how

can we use it (and variants, like 'quus') at the same time? So I suppose that the sceptic assumes that he and I agree in our *present* uses of the word "plus": we both use it to denote addition. He does *not*—at least initially—deny or doubt that addition is a genuine function, defined on all pairs of integers, nor does he deny that we can speak of it. Rather he asks why I now believe that by "plus" in the *past,* I meant addition rather than quaddition. If I meant the former, then to accord with my previous usage I should say '125' when asked to give the result of calculating '68 plus 57'. If I meant the latter, I should say '5'.

The present exposition tends to differ from Wittgenstein's original formulations in taking somewhat greater care to make explicit a distinction between use and mention, and between questions about present and past usage. About the present example Wittgenstein might simply ask, "How do I know that I should respond '125' to the query '68 + 57'?" or "How do I know that '68 + 57' comes out 125?" I have found that when the problem is formulated this way, some listeners hear it as a sceptical problem about *arithmetic:* "How do I know that 68 + 57 is 125?" (Why not answer this question with a mathematical proof?) At least at this stage, scepticism about arithmetic should not be taken to be in question: we may assume, if we wish, that 68 + 57 *is* 125. Even if the question is reformulated 'metalinguistically' as "How do I know that 'plus', as I use it, denotes a function that, when applied to 68 and 57, yields 125?", one may answer, "Surely I know that 'plus' denotes the plus function and accordingly that '68 plus 57' denotes 68 plus 57. But if I know arithmetic, I know that 68 plus 57 is 125. So I know that '68 plus 57' denotes 125!" And surely, if I use language at all, I cannot doubt coherently that "plus," as I now use it, denotes plus! Perhaps I cannot (at least at this stage) doubt this about my *present* usage. But I can doubt that my *past* usage of "plus" denoted plus. The previous remarks—about a frenzy and LSD—should make this quite clear.

Let me repeat the problem. The sceptic doubts whether any instructions I gave myself in the past compel (or justify) the answer '125' rather than '5'. He puts the challenge in terms of a sceptical hypothesis about a change in my usage. Perhaps when I used the term "plus" in the *past,* I always meant quus: by hypothesis I never gave myself any explicit directions that were incompatible with such a supposition.

Of course, ultimately, if the sceptic is right, the concepts of meaning and of intending one function rather than another will make no sense. For the sceptic holds that no fact about my past history—nothing that was ever in my mind, or in my external behavior—establishes that I meant plus rather than quus. (Nor, of course, does any fact establish that I meant quus!) But if this is correct, there can of course be no fact about which function I meant, and if there can be no fact about which particular function I meant in the *past,* there can be none in the *present* either. But before we pull the rug out from under our own feet, we begin by speaking as if the notion that at present we mean a certain function by "plus" is unquestioned and unquestionable. Only *past* usages are to be questioned. Otherwise, we will be unable to *formulate* our problem.

Another important rule of the game is that there are no limitations, in particular, no *behaviorist* limitations, on the facts that may be cited to answer the sceptic. The evidence is not to be confined to that available to an external observer, who can observe my overt behavior but not my internal mental state. It would be interesting if nothing in my external behavior could show whether I meant plus or quus, but something about my inner state could. But the problem here is more radical. Wittgenstein's philosophy of mind has often been viewed as behavioristic, but to the extent that Wittgenstein may (or may not) be hostile to the 'inner', no such hostility is to be assumed as a premise; it is to be argued as a conclusion. So whatever 'looking into my mind' may be, the sceptic asserts that even if God were to do it, he still could not determine that I meant addition by "plus."

This feature of Wittgenstein contrasts, for example, with Quine's discussion of the 'indeterminacy of translation'.[6] There are many points of contact between Quine's discussion

and Wittgenstein's. Quine, however, is more than content to assume that only behavioral evidence is to be admitted into his discussion. Wittgenstein, by contrast, undertakes an extensive introspective[7] investigation, and the results of the investigation, as we shall see, form a key feature of his argument. Further, the way the sceptical doubt is presented is not behavioristic. It is presented from the 'inside'. Whereas Quine presents the problem about meaning in terms of a linguist, trying to guess what someone *else* means by his words on the basis of his behavior, Wittgenstein's challenge can be presented to me as a question about *myself:* was there some past fact about me—what I 'meant' by plus—that mandates what I should do now?

To return to the sceptic. The sceptic argues that when I answered '125' to the problem '68 + 57', my answer was an unjustified leap in the dark; my past mental history is equally compatible with the hypothesis that I meant quus, and therefore should have said '5'. We can put the problem this way: When asked for the answer to '68 + 57', I unhesitatingly and automatically produced '125', but it would seem that if previously I never performed this computation explicitly I might just as well have answered '5'. Nothing justifies a brute inclination to answer one way rather than another. . . . Wittgenstein has invented a new form of scepticism. Personally I am inclined to regard it as the most radical and original sceptical problem that philosophy has seen to date, one that only a highly unusual cast of mind could have produced. Of course he does not wish to leave us with his problem, but to solve it: the sceptical conclusion is insane and intolerable. It is his solution, I will argue, that contains the argument against 'private language'; for allegedly, the solution will not admit such a language. But it is important to see that his achievement in posing this problem stands on its own, independently of the value of his own solution of it and the resultant argument against private language. For, if we see Wittgenstein's problem as a real one, it is clear that he has often been read from the wrong perspective. Readers, my previous self certainly included, have often been inclined to wonder: "How can he prove private language impossible? How can I possibly have any difficulty identifying my own sensations? And if there were a difficulty, how could 'public' criteria help me? I must be in pretty bad shape if I needed external *help* to identify my own sensations!" But if I am right, a proper orientation would be the opposite. The main problem is *not,* "How can we show private language—or some other special form of language—to be *impossible?*"; rather it is, "How can we show *any language* at all (public, private, or what-have-you) to be *possible?*"[8] It is not that calling a sensation 'pain' is easy, and Wittgenstein must invent a difficulty.[9] On the contrary, Wittgenstein's main problem is that it appears that he has shown *all* language, *all* concept formation, to be impossible, indeed unintelligible.

It is important and illuminating to compare Wittgenstein's new form of scepticism with the classical scepticism of Hume; there are important analogies between the two. Both develop a sceptical paradox, based on questioning a certain *nexus* from past to future. Wittgenstein questions the nexus between past 'intention' or 'meanings' and present practice: for example, between my past 'intentions' with regard to 'plus' and my present computation '68 + 57 = 125'. Hume questions two other nexuses, related to each other: the causal nexus whereby a past event necessitates a future one, and the inductive inferential nexus from the past to the future.

The analogy is obvious. It has been obscured for several reasons. First, the Humean and the Wittgensteinian problems are of course distinct and independent, though analogous. Second, Wittgenstein shows little interest in or sympathy with Hume: he has been quoted as saying that he could not read Hume because he found it "a torture".[10] Furthermore, Hume is the prime source of some ideas on the nature of mental states that Wittgenstein is most concerned to attack.[11] Finally (and probably most important), Wittgenstein never avows, and almost surely would not avow, the label 'sceptic', as Hume explicitly did. Indeed, he has often appeared to be a 'common-sense' philosopher, anxious to defend our ordinary

conceptions and dissolve traditional philosophical doubts. Is it not Wittgenstein who held that philosophy only states what everyone admits?

Yet even here the difference between Wittgenstein and Hume should not be exaggerated. Even Hume has an important strain, dominant in some of his moods, that the philosopher never questions ordinary beliefs. Asked whether he "be really one of those sceptics, who hold that all is uncertain," Hume replies "that this question is entirely superfluous, and that neither I, nor any other person, was ever sincerely and constantly of that opinion."[12] Even more forcefully, discussing the problem of the external world: "We may well ask, *What causes induce us to believe in the existence of body?* but 'tis in vain to ask, *Whether there be body or not?* That is a point, which we must take for granted in all our reasonings."[13] Yet this oath of fealty to common sense begins a section that otherwise looks like an argument that the common conception of material objects is irreparably incoherent!

When Hume is in a mood to respect his professed determination never to deny or doubt our common beliefs, in what does his 'scepticism' consist? First, in a sceptical *account* of the causes of these beliefs; and second, in sceptical analyses of our common notions. In some ways Berkeley, who did not regard his own views as sceptical, may offer an even better analogy to Wittgenstein. At first blush, Berkeley, with his denial of matter, and of any objects 'outside the mind' seems to be *denying* our common beliefs; and for many of us the impression persists through later blushes. But not for Berkeley. For him, the impression that the common man is committed to matter and to objects outside the mind derives from an erroneous metaphysical interpretation of common talk. When the common man speaks of an 'external material object' he does not really mean (as we might say *sotto voce*) an *external material object* but rather he means something like 'an idea produced in me independently of my will'.[14]

Berkeley's stance is not uncommon in philosophy. The philosopher advocates a view apparently in patent contradiction to common sense. Rather than repudiating common sense, he asserts that the conflict comes from a philosophical misinterpretation of common language—sometimes he adds that the misinterpretation is encouraged by the 'superficial form' of ordinary speech. He offers his own analysis of the relevant common assertions, one that shows that they do not really say what they seem to say. For Berkeley this philosophical strategy is central to his work. To the extent that Hume claims that he merely analyses common sense and does not oppose it, he invokes the same strategy as well. The practice can hardly be said to have ceased today.[15]

Personally I think such philosophical claims are almost invariably suspect. What the claimant calls a 'misleading philosophical misconstrual' of the ordinary statement is probably the natural and correct understanding. The real misconstrual comes when the claimant continues, "All the ordinary man really means is . . ." and gives a sophisticated analysis compatible with his own philosophy. Be this as it may, the important point for present purposes is that Wittgenstein makes a Berkeleyan claim of this kind. For—as we shall see—his solution to his own sceptical problem begins by agreeing with the sceptics that there is no 'superlative fact' (§192) about my mind that constitutes my meaning addition by "plus" and determines in advance what I should do to accord with this meaning. But, he claims (in §§183–93), the appearance that our ordinary concept of meaning demands such a fact is based on a philosophical misconstrual—albeit a natural one—of such ordinary expressions as "he meant such-and-such," "the steps are determined by the formula," and the like. How Wittgenstein construes these expressions we shall see presently. For the moment let us only remark that Wittgenstein thinks that any construal that looks for something in my present mental state to differentiate between my meaning addition or quaddition, or that will consequently show that in the future I should say '125' when asked about '68 + 57', *is* a misconstrual and attributes to the ordinary man a notion of meaning that *is* refuted by the sceptical argument. "We are," he says in

§194—note that Berkeley could have said just the same thing!—"like savages, primitive people, who hear the expressions of civilized men, put a false interpretation on them, and then draw the queerest conclusions from it." Maybe so. Personally I can only report that, in spite of Wittgenstein's assurances, the 'primitive' interpretation often sounds rather good to me . . .

In his *Enquiry,* after he has developed his "Sceptical Doubts Concerning the Operations of the Understanding," Hume gives his "Sceptical Solution of These Doubts." What is a 'sceptical' solution? Call a proposed solution to a sceptical philosophical problem a *straight* solution if it shows that on closer examination the scepticism proves to be unwarranted; an elusive or complex argument proves the thesis the sceptic doubted. Descartes gave a 'straight' solution in this sense to his own philosophical doubts. An a priori justification of inductive reasoning, and an analysis of the causal relation as a genuine necessary connection or nexus between pairs of events, would be straight solutions of Hume's problems of induction and causation, respectively. A *sceptical* solution of a sceptical philosophical problem begins on the contrary by conceding that the sceptic's negative assertions are unanswerable. Nevertheless our ordinary practice or belief is justified because—contrary appearances notwithstanding—it need not require the justification the sceptic has shown to be untenable. And much of the value of the sceptical argument consists precisely in the fact that he has shown that an ordinary practice, if it is to be defended at all, cannot be defended in a certain way. A sceptical solution may also involve—in the manner suggested above—a sceptical analysis or account of ordinary beliefs to rebut their *prima facie* reference to a metaphysical absurdity.

The rough outlines of Hume's sceptical solution to his problem are well known.[16] Not an a priori argument, but custom, is the source of our inductive inferences. If A and B are two types of events which we have seen constantly conjoined, then we are conditioned—Hume is a grandfather of this modern psychological notion—to expect an event of type B on being presented with one of type A. To say of a particular event a that it caused another event b is to place these two events under two types, A and B, which we expect to be constantly conjoined in the future as they were in the past. The idea of necessary connection comes from the 'feeling of customary transition' between our ideas of these event types.

The philosophical merits of the Humean solution are not our present concern. Our purpose is to use the analogy with the Humean solution to illuminate Wittgenstein's solution to his own problem. For comparative purposes one further consequence of Hume's sceptical solution should be noted. Naively, one might suppose that whether a particular event a causes another particular event b, is an issue solely involving the events a and b alone (and their relations), and involves no other events. If Hume is right, this is not so. Even if God were to look at the events, he would discern nothing relating them other than that one succeeds the other. Only when the particular events a and b are thought of as subsumed under two respective event types, A and B, which are related by a generalization that *all* events of type A are followed by events of type B, can a be said to 'cause' b. When the events a and b are considered by themselves alone, no causal notions are applicable. This Humean conclusion might be called: the impossibility of private causation.

Can one reasonably protest: surely there is nothing the event a can do with the *help* of other events of the same type that it cannot do by itself! Indeed, to say that a, by itself, is a sufficient cause of b is to say that, had the rest of the universe been removed, a still would have produced b! Intuitively this may well be so, but the intuitive objection ignores Hume's sceptical argument. The whole point of the sceptical argument is that the common notion of one event 'producing' another, on which the objection relies, is in jeopardy. It appears that there is no such relation as 'production' at all, that the causal relation is fictive. After the sceptical argument has been seen to be unanswerable on its own terms, a sceptical solution is offered, containing all we can salvage of the notion of causation. It just is a feature of this

analysis that causation makes no sense when applied to two isolated events, with the rest of the universe removed. Only inasmuch as these events are thought of as instances of event types related by a regularity can they be thought of as causally connected. If two particular events were somehow so *sui generis* that it was logically excluded that they be placed under any (plausibly natural) event types, causal notions would not be applicable to them.

Of course I am suggesting that Wittgenstein's argument against private language has a structure similar to Hume's argument against private causation. Wittgenstein also states a sceptical paradox. Like Hume, he accepts his own sceptical argument and offers a 'sceptical solution' to overcome the appearance of paradox. His solution involves a sceptical interpretation of what is involved in such ordinary assertions as "Jones means addition by '+'." The impossibility of private language emerges as a corollary of his sceptical solution of his own paradox, as does the impossibility of 'private causation' in Hume. It turns out that the sceptical solution does not allow us to speak of a single individual, considered by himself and in isolation, as ever meaning anything. Once again an objection based on an intuitive feeling that no one else can affect what I mean by a given symbol ignores the sceptical argument that undermines any such naive intuition about meaning.

I have said that Wittgenstein's solution to his problem is a sceptical one. He does not give a 'straight' solution, pointing out to the silly sceptic a hidden fact he overlooked, a condition in the world which constitutes my meaning addition by "plus." In fact, he agrees with his own hypothetical sceptic that there is no such fact, no such condition in either the 'internal' or the 'external' world. Admittedly, I am expressing Wittgenstein's view more straightforwardly than he would ordinarily allow himself to do. For in denying that there is any such fact, might we not be expressing a philosophical thesis that doubts or denies something everyone admits? We do not wish to doubt or deny that when people speak of themselves and others as meaning something by their words, as following rules, they do so with perfect right. We do not even wish to deny the propriety of an ordinary use of the phrase "the fact that Jones meant addition by such-and-such a symbol," and indeed such expressions do have perfectly ordinary uses. We merely wish to deny the existence of the 'superlative fact' that philosophers misleadingly attach to such ordinary forms of words, not the propriety of the forms of words themselves.

It is for this reason that I conjectured above that Wittgenstein's professed inability to write a work with conventionally organized arguments and conclusions stems at least in part, not from personal and stylistic proclivities, but from the nature of his work. Had Wittgenstein—contrary to his notorious and cryptic maxim in §128—stated the outcomes of his conclusions in the form of definite theses, it would have been very difficult to avoid formulating his doctrines in a form that consists in apparent sceptical denials of our ordinary assertions. Berkeley runs into similar difficulties. Partly he avoids them by stating his thesis as the denial of the existence of 'matter', and claiming that 'matter' is a bit of philosophical jargon, not expressive of our common-sense view. Nevertheless he is forced at one point to say—apparently contrary to his usual official doctrine—that he denies a doctrine "strangely prevailing amongst men."[17] If, on the other hand, we do not state our conclusions in the form of broad philosophical theses, it is easier to avoid the danger of a denial of any ordinary belief, even if our imaginary interlocuter (e.g. §189; see also §195)[18] accuses us of doing so. Whenever our opponent insists on the perfect propriety of an ordinary form of expression (e.g. that "the steps are determined by the formula," "the future application is already present"), we can insist that if these expressions are properly understood, we agree. The danger comes when we try to give a precise formulation of exactly what it is that we *are* denying—*what* 'erroneous interpretation' our opponent is placing on ordinary means of expression. It may be hard to do this without producing yet an-

other statement that, we must admit, is *still* 'perfectly all right, properly understood'.

So Wittgenstein, perhaps cagily, might well disapprove of the straightforward formulation given here. Nevertheless I choose to be so bold as to say: Wittgenstein holds, with the sceptic, that there is no fact as to whether I mean plus or quus. . . . Let me, then, summarize the 'private language argument' as it is presented in this essay. (1) We all suppose that our language expresses concepts—"pain," "plus," "red"—in such a way that, once I 'grasp' the concept, all future applications of it are determined (in the sense of being uniquely *justified* by the concept grasped). In fact, it seems that no matter what is in my mind at a given time, I am free in the future to interpret it in different ways—for example, I could follow the sceptic and interpret "plus" as "quus." In particular, this point applies if I direct my attention to a sensation and name it; nothing I have done determines future applications (in the justificatory sense above). Wittgenstein's scepticism about the determination of future usage by the past contents of my mind is analogous to Hume's scepticism about the determination of the future by the past (causally and inferentially). (2) The paradox can be resolved only by a 'sceptical solution of these doubts', in Hume's classic sense. This means that we must give up the attempt to find any fact about me in virtue of which I mean "plus" rather than "quus," and must then go on in a certain way. Instead we must consider how we actually use: (i) the categorical assertion that an individual is following a given rule (that he means addition by 'plus'); (ii) the conditional assertion that "if an individual follows such-and-such a rule, he must do so-and-so on a given occasion" (e.g., "if he means addition by '+', his answer to '68 + 57' should be '125' "). That is to say, we must look at the circumstances under which these assertions are introduced into discourse, and their role and utility in our lives. (3) As long as we consider a single individual in isolation, all we can say is this: An individual often does have the experience of being confident that he has 'got' a certain rule (sometimes that he has grasped it "in a flash"). It is an empirical fact that, after that experience, individuals often are disposed to give responses in concrete cases with complete confidence that proceeding this way is 'what was intended'. We cannot, however, get any further in explaining on this basis the use of the conditionals in (ii) above. Of course, dispositionally speaking, the subject is indeed determined to respond in a certain way, say, to a given addition problem. Such a disposition, together with the appropriate 'feeling of confidence', could be present, however, even if he were not really following a rule at all, or even if he were doing the 'wrong' thing. The justificatory element of our use of conditionals such as (ii) is unexplained. (4) If we take into account the fact that the individual is in a community, the picture changes and the role of (i) and (ii) above becomes apparent. When the community accepts a particular conditional (ii), it accepts its *contraposed* form: the failure of an individual to come up with the particular responses the community regards as right leads the community to suppose that he is not following the rule. On the other hand, if an individual passes enough tests, the community (endorsing assertions of the form (i)) accepts him as a rule follower, thus enabling him to engage in certain types of interactions with them that depend on their reliance on his responses. Note that this solution explains how the assertions in (i) and (ii) are introduced into language; it does *not* give conditions for these statements to be true. (5) The success of the practices in (3) depends on the brute empirical fact that we agree with each other in our responses. Given the sceptical argument in (1), this success cannot be explained by 'the fact that we all grasp the same concepts'. (6) Just as Hume thought he had demonstrated that the causal relation between two events is unintelligible unless they are subsumed under a regularity, so Wittgenstein thought that the considerations in (2) and (3) above showed that all talk of an individual following rules has reference to him as a member of a community, as in (3). In particular, for the

conditionals of type (ii) to make sense, the community must be able to judge whether an individual is indeed following a given rule in particular applications, i.e. whether his responses agree with their own. In the case of avowals of sensations, the way the community makes this judgement is by observing the individual's behavior and surrounding circumstances.

A few concluding points regarding the argument ought to be noted. First, following §243, a 'private language' is usually defined as a language that is logically impossible for anyone else to understand. The private language argument is taken to argue against the possibility of a private language in this sense. This conception is not in error, but it seems to me that the emphasis is somewhat misplaced. What is really denied is what might be called the 'private model' of rule following, that the notion of a person following a given rule is to be analyzed simply in terms of facts about the rule follower and the rule follower alone, without reference to his membership in a wider community. (In the same way, what Hume denies is the private model of causation: that whether one event causes another is a matter of the relation between these two events alone, without reference to their subsumption under larger event types.) The impossibility of a private language in the sense just defined does indeed follow from the incorrectness of the private model for language and rules, since the rule following in a 'private language' could only be analyzed by a private model, but the incorrectness of the private model is more basic, since it applies to all rules. I take all this to be the point of §202.

Does this mean that Robinson Crusoe, isolated on an island, cannot be said to follow any rules, no matter what he does?[19] I do not see that this follows. What does follow is that *if* we think of Crusoe as following rules, we are taking him into our community and applying our criteria for rule following to him.[20] The falsity of the private model need not mean that a *physically isolated* individual cannot be said to follow rules; rather that an individual, *considered in isolation* (whether or not he is physically isolated), cannot be said to do so. Remember that Wittgenstein's theory is one of assertability conditions. Our community can assert of any individual that he follows a rule if he passes the tests for rule following applied to any member of the community.

Finally, the point just made in the last paragraph, that Wittgenstein's theory is one of assertability conditions, deserves emphasis. Wittgenstein's theory should not be confused with a theory that, for any *m* and *n,* the value of the function we mean by "plus," *is* (by definition) the value that (nearly) all the linguistic community would give as the answer. Such a theory would be a theory of the *truth* conditions of such assertions as "By 'plus' we mean such-and-such a function," or "By 'plus' we mean a function, which, when applied to 68 and 57 as arguments, yields 125 as value." (An infinite, exhaustive totality of specific conditions of the second form would determine which function was meant, and hence would determine a condition of the first form.) The theory would assert that 125 is the value of the function meant for given arguments, if and only if '125' is the response nearly everyone would give, given these arguments. Thus the theory would be a social, or community-wide, version of the dispositional theory, and would be open to at least some of the same criticisms as the original form. I take Wittgenstein to deny that he holds such a view, for example, in *Remarks on the Foundations of Mathematics,* v, §33 [vii, §40]: "Does this mean, e.g., that the definition of the same would be this: same is what all or most human beings take for the same?—Of course not."[21] (See also *Philosophical Investigations,* p. 226, "Certainly the propositions, 'Human beings believe that twice two is four' and 'Twice two is four' do not mean the same"; and see also §§240–1.) One must bear firmly in mind that Wittgenstein has no theory of truth-conditions—necessary and sufficient conditions—for the correctness of one response rather than another to a new addition problem. Rather he simply points out that each of us *automatically* calculates new addition problems (without feeling the need to check with the community whether our procedure is proper); that the

community feels entitled to correct a deviant calculation; that in practice such deviation is rare, and so on. Wittgenstein thinks that these observations about sufficient conditions for justified assertion are enough to illuminate the role and utility in our lives of assertion about meaning and determination of new answers. What follows from these assertability conditions is *not* that the answer everyone gives to an addition problem is, by definition, the correct one, but rather the platitude that, if everyone agrees upon a certain answer, then no one will feel justified in calling the answer wrong.

Obviously there are countless relevant aspects of Wittgenstein's philosophy of mind that I have not discussed.[22] About some aspects I am not clear, and others have been left untouched because of the limits of this essay.[23] In particular, I have not discussed numerous issues arising out of the paragraphs *following* §243 that are usually called the 'private language argument', nor have I really discussed Wittgenstein's attendant positive account of the nature of sensation language and of the attribution of psychological states. Nevertheless, I do think that the basic 'private language argument' precedes these passages, and that only with an understanding of this argument can we begin to comprehend or consider what follows. That was the task undertaken in this essay.

NOTES

1. Unless otherwise specified (explicitly or contextually), references are to *Philosophical Investigations*. The small numbered units of the *Investigations* are termed 'sections' (or 'paragraphs'). Page references are used only if a section reference is not possible, as in the second part of the *Investigations*. Throughout I quote the standard printed English translation (by G. E. M. Anscombe) and make no attempt to question it except in a very few instances. *Philosophical Investigations* has undergone several editions since its first publication in 1953 but the paragraphing and pagination remain the same. The publishers are Basil Blackwell, Oxford and Macmillan, New York.

   This essay does not proceed by giving detailed exegesis of Wittgenstein's text but rather develops the arguments in its own way. I recommend that the reader reread the *Investigations* in the light of the present exegesis and see whether it illuminates the text.
2. Basil Blackwell, Oxford: 1956. In the first edition of *Remarks on the Foundation of Mathematics* the editors assert (p. vi) that Wittgenstein appears originally to have intended to include some of the material on mathematics in *Philosophical Investigations*. The third edition (1978) includes more material than earlier editions and rearranges some of the sections and divisions of earlier editions. When I wrote the present work, I used the first edition. Where the references differ, the equivalent third edition reference is given in square brackets.
3. Personally I feel, however, that the role of stylistic considerations here cannot be denied. It is clear that purely stylistic and literary considerations meant a great deal to Wittgenstein. His own stylistic preference obviously contributes to the difficulty of his work as well as to its beauty.
4. Perhaps I should make a remark about such expressions as "By 'plus' I meant quus (or plus)," "By 'green' I meant green," etc. I am not familiar with an accepted felicitous convention to indicate the object of the verb 'to mean'. There are two problems. First, if one says, "By 'the woman who discovered radium' I meant the woman who discovered radium," the object can be interpreted in two ways. It may stand for a woman (Marie Curie), in which case the assertion is true only if 'meant' is used to mean referred to (as it can be used); or it may be used to denote the *meaning* of the quoted expression, not a woman, in which case the assertion is true with 'meant' used in the ordinary sense. Second, as is illustrated by 'referred to', 'green', 'quus', etc. above, as objects of 'meant', one must use various expressions as objects in an awkward manner contrary to normal grammar. (Frege's difficulties concerning unsaturatedness are related.) Both problems tempt one to put the object in quotation marks, like the subject; but such a usage conflicts with the convention of philosophical logic that a quotation denotes the expression quoted. Some special 'meaning marks', as proposed for example by David Kaplan, could be useful here. If one is content to ignore the first difficulty and always use 'mean' to mean denote (for most purposes of the present paper, such a reading would suit at least as well as an intensional one; often I speak as if it is a *numerical function* that is meant by plus), the second problem might lead one to nominalize the objects—'plus' denotes the plus function, 'green' denotes greeness, etc. I contemplated using italics (" 'plus' means *plus*"; " 'mean' may mean *denote*"), but I decided that normally (except when italics are otherwise appropriate, especially when a neologism like 'quus' is intro-

duced for the first time), I will write the object of 'to mean' as an ordinary roman object. The convention I have adopted reads awkwardly in the written language but sounds rather reasonable in the spoken language.

Since use–mention distinctions are significant for the argument as I give it, I try to remember to use quotation marks when an expression is mentioned. However, quotation marks are also used for other purposes where they might be invoked in normal non-philosophical English writing (for example, in the case of " 'meaning marks' " in the previous paragraph, or " 'quasi-quotation' " in the next sentence). Readers familiar with Quine's 'quasi-quotation' will be aware that in some cases I use ordinary quotation where logical purity would require that I use quasi-quotation or some similar device. I have not tried to be careful about this matter, since I am confident that in practice readers will not be confused.

5. I believe I got the phrase "both inside and outside language" from a conversation with Rogers Albritton.
6. See W. V. Quine, *Word and Object* (MIT, The Technology Press, Cambridge, Massachusetts: 1960) especially chapter 2, "Translation and Meaning" (pp. 26–79). See also *Ontological Relativity and Other Essays* (Columbia University Press, New York and London: 1969), especially the first three chapters (pp. 1–90); and see also "On the Reasons for the Indeterminacy of Translation," *The Journal of Philosophy,* vol. 67 (1970), pp. 178–83.
7. I do not mean the term 'introspective' to be laden with philosophical doctrine. Of course much of the baggage that has accompanied this term would be objectionable to Wittgenstein in particular. I simply mean that he makes use, in his discussion, of our own memories and knowledge of our 'inner' experiences.
8. So put, the problem has an obvious Kantian flavor.
9. See especially the discussions of 'green' and 'grue' above, [not reprinted in this volume] which plainly could carry over to pain (let 'pickle' apply to pains before *t,* and tickles thereafter!); but it is clear enough by now that the problem is completely general.
10. Karl Britton, "Portrait of a Philosopher," *The Listener,* LIII, no. 1372 (June 16, 1955), p. 1072, quoted by George Pitcher, *The Philosophy of Wittgenstein* (Prentice Hall, Englewood Cliffs, N.J.: 1964), p. 325.
11. Much of Wittgenstein's argument can be regarded as an attack on characteristically Humean (or classical empiricist) ideas. Hume posits an introspectible qualitative state for each of our psychological states (an 'impression'). Further, he thinks that an appropriate 'impression' or 'image' can constitute an 'idea', without realizing that an image in no way tells us how it is to be applied. Of course the Wittgensteinian paradox is, among other things, a strong protest against such suppositions.
12. David Hume, *A Treatise of Human Nature,* ed. L. A. Selby-Bigge, Clarendon Press, Oxford: 1888), Book I, part IV, section I (p. 183 in the Selby-Bigge edition).
13. Hume, *ibid.,* Book I, part IV, Section II (p. 187 in the Selby-Bigge edition). Hume's occasional affinities to 'ordinary language' philosophy should not be overlooked. Consider the following: "Those philosophers, who have divided human reason into *knowledge and probability,* and have defined the first to be *that evidence, which arises from the comparison of ideas,* are obliged to comprehend all our arguments from causes or effects under the general term of probability. But tho' everyone be free to use his terms in what sense he pleases . . . 'tis however certain, that in common discourse we readily affirm, that many arguments from causation exceed probability, and may be received as a superior kind of evidence. One would appear ridiculous, who would say, that 'tis only probable the sun will rise tomorrow, or that all men must dye . . ." (*ibid.,* Book I, part III, section XI, p. 124 in the Selby-Bigge edition).
14. George Berkeley, *The Principles of Human Knowledge,* §§29–34. Of course the characterization may be oversimplified, but it suffices for present purposes.
15. It is almost 'analytic' that I cannot produce a common contemporary example that would not meet with vigorous opposition. Those who hold the cited view would argue that, in this case, their analyses of ordinary usage are really correct. I have no desire to enter into an irrelevant controversy here, but I myself find that many of the 'topic-neutral' analyses of discourse about the mind proposed by contemporary materialists are just the other side of the Berkeleyan coin.
16. Writing this sentence, I find myself prey to an appropriate fear that (some) experts in Hume and Berkeley will not approve of some particular thing that I say about these philosophers here. I have made no careful study of them for the purpose of this paper. Rather a crude and fairly conventional account of the 'rough outlines' of their views is used for purposes of comparison with Wittgenstein.
17. Berkeley, *The Principles of Human Knowledge,* §4. Of course Berkeley might mean that the prevalence of the doctrine stems from the influence of philosophical theory rather than common sense, as indeed he asserts in the next section.
18. §189: "But *are* the steps then *not* determined by the algebraic formula?" In spite of Wittgenstein's interpretation within his own philosophy

of the ordinary phrase "the steps are determined by the formula", the impression persists that the interlocutor's characterization of his view is really correct. See §195: "But I don't mean that what I do now (in grasping a sense) determines the future use *causally* and as a matter of experience, but that in a *queer* way, the use itself is in some sense present," which are the words of the interlocutor, and the bland reply. "But of course it is, 'in *some* sense'! Really the only thing wrong with what you say is the expression "in a queer way". The rest is all right; and the sentence only seems queer when one imagines a different language-game for it from the one in which we actually use it."

19. See . . . A. J. Ayer, "Can there be a Private Language?" *Proceedings of the Aristotelian Society Supplementary Volume* 28 (1954). Ayer assumes that the 'private language argument' excludes Crusoe from language [and] takes this alleged fact to be fatal to Wittgenstein's argument. . . . Others, pointing out that a 'private language' is one that others *cannot* understand (see the preceding paragraph in the text), see no reason to think that the 'private language argument' has anything to do with Crusoe (as long as we could understand his language). My own view of the matter, as explained very briefly in the text, differs somewhat from all these opinions.
20. If Wittgenstein would have any problem with Crusoe, perhaps the problem would be whether we have any 'right' to take him into our community in this way, and attribute our rules to him. See Wittgenstein's discussion of a somewhat similar question in §§199–200, and his conclusion, "Should we still be inclined to say they were playing a game? What right would one have to say so?"
21. Although, in the passage in question, Wittgenstein is speaking of a particular language-game of bringing something else and bringing the same, it is clear in context that it is meant to illustrate his general problem about rules. The entire passage is worth reading for the present issue.
22. [. . .] As members of the community correct each other, might a given individual correct himself? Some question such as this was prominent in earlier discussions of verificationist versions of the private language argument. Indeed, in the absence of Wittgenstein's sceptical paradox, it would appear that an individual remembers his own 'intentions' and can use one memory of these intentions to correct another mistaken memory. In the presence of the paradox, any such 'naive' ideas are meaningless. Ultimately, an individual may simply have conflicting brute inclinations, while the upshot of the matter depends on his will alone. The situation is not analogous to the case of the community, where distinct individuals have distinct and independent wills, and where, when an individual is accepted into the community, others judge that they can rely on his response (as was described in the text above). No corresponding relation between an individual and himself has the same utility. Wittgenstein may be indicating something like this in §268.
23. I might mention that, in addition to the Humean analogy emphasized in this essay, it has struck me that there is perhaps a certain analogy between Wittgenstein's private language argument and Ludwig von Mises's celebrated argument concerning economic calculation under socialism. (See e.g., his *Human Action,* 2d ed., Yale University Press, New Haven: 1963, chapter 26, pp. 698–715, for one statement.) According to Mises, a rational economic calculator (say, the manager of an industrial plant) who wishes to choose the most efficient means to achieve given ends must compare alternative courses of action for cost effectiveness. To do this, he needs an array of prices (e.g. of raw materials, or machinery) set by *others.* If *one* agency set *all* prices, it could have no rational basis to choose between alternative courses of action. (Whatever seemed to it to be right would be right, so one cannot talk about right.) I do not know whether the fact bodes at all ill for the private language argument, but my impression is that although it is usually acknowledged that Mises's argument points to a real difficulty for centrally planned economies, it is now almost universally rejected as a theoretical proposition.