

## The Solution and the 'Private Language' Argument

The sceptical argument, then, remains unanswered. There can be no such thing as meaning anything by any word. Each new application we make is a leap in the dark; any present intention could be interpreted so as to accord with anything we may choose to do. So there can be neither accord, nor conflict. This is what Wittgenstein said in §202.

Wittgenstein's sceptical problem is related to some work of two other recent writers who show little direct influence from Wittgenstein. Both have already been mentioned above. The first is W. V. Quine,<sup>37</sup> whose well-known theses of the indeterminacy of translation and the inscrutability of reference also question whether there are any objective facts as to what we mean. If I may anticipate matters that the present exposition has not yet introduced, Quine's emphasis on agreement is obviously congenial to Wittgenstein's view.<sup>38</sup> So

<sup>37</sup> See pp. 14–15 above, and note 10.

<sup>38</sup> For 'agreement' and the related notion of 'form of life' in Wittgenstein, see pp. 96–8 below. Quine, *Word and Object*, p. 27, characterizes language as "the complex of present dispositions to verbal behavior, in which speakers of the same language have perforce come to resemble one another"; also, see §2, *Word and Object*, pp. 5–8. Some of the major

is his rejection of any notion that inner 'ideas' or 'meanings' guide our linguistic behavior. However, there are differences. As I have remarked above, Quine bases his argument from the outset on behavioristic premises. He would never emphasize introspective thought experiments in the way Wittgenstein does, and he does not think of views that posit a private inner world as in need of elaborate refutation. For Quine, the untenability of any such views should be obvious to anyone who accepts a modern scientific outlook. Further, since Quine sees the philosophy of language within a hypothetical framework of behavioristic psychology, he thinks of problems about meaning as problems of disposition to behavior. This orientation seems to have consequences for the form of Quine's problem as opposed to Wittgenstein's. The important problem for Wittgenstein is that my present mental state does not appear to determine what I *ought* to do in the future. Although I may *feel* (now) that something in my head corresponding to the word 'plus' mandates a determinate response to any new pair of arguments, in fact nothing in my head does so. Alluding to one of Wittgenstein's earliest examples, 'ostensive' learning of the color word 'sepia' (§§28-30),<sup>39</sup> Quine protests against Wittgenstein that, given our 'inborn propensity to find one stimulation qualitatively more akin to a second stimulation than to a third' and sufficient conditioning 'to eliminate wrong generalizations', eventually the term will be learnt: ". . . in principle nothing more is needed in learning 'sepia' than in any conditioning or induction."<sup>40</sup> By "learning 'sepia'", Quine means developing the right disposition to apply 'sepia' in particular cases. It should be clear from Wittgenstein's text that he too is aware, indeed emphasizes, that in practice there need be no difficulty

concepts of *Word and Object*, such as that of 'observation sentence', depend on this uniformity in the community. Nevertheless, agreement seems to have a more crucial role in Wittgenstein's philosophy than in Quine's.

<sup>39</sup> This example is discussed below. See pp. 83-4 and note 72.

<sup>40</sup> Quine, *Ontological Relativity and Other Essays*, p. 31.

in this sense about the learning of 'sepia'. The fundamental problem, as I have stated it earlier, is different: whether my actual dispositions are 'right' or not, is there anything that mandates what they *ought* to be? Since Quine formulates the issues dispositionally, this problem cannot be stated within his framework. For Quine, since any fact as to whether I mean plus or quus will show up in my behavior, there is no question, given my disposition, as to what I mean.

It has already been argued above that such a formulation of the issues seems inadequate. My actual dispositions are not infallible, nor do they cover all of the infinitely many cases of the addition table. However, since Quine does see the issues in terms of dispositions, he is concerned to show that even if dispositions were ideally seen as infallible and covering all cases, there are still questions of interpretation that are left undetermined. First, he argues roughly that the interpretation of sufficiently 'theoretical' utterances, not direct observation reports, is undetermined even by all my ideal dispositions. Further, he seeks to show by examples such as 'rabbit' and 'rabbit-stage' that, even given fixed interpretation of our sentences as wholes and certainly given all our ideal dispositions to behavior, the interpretation (reference) of various lexical items is still not fixed.<sup>41</sup> These are interesting claims, distinct from Wittgenstein's. For those of us who are not as behavioristically inclined as Quine, Wittgenstein's problem may lead to a new look at Quine's theses. Given Quine's own formulation of his theses, it appears open to a non-behaviorist to regard his arguments, *if* he accepts them, as demonstrations that any behavioristic account of meaning must be inadequate — it cannot even distinguish between a word meaning rabbit and one meaning rabbit-stage. But if Wittgenstein is right, and no amount of access to my mind can reveal whether I mean plus or quus, may the same not hold for rabbit and rabbit-stage? So perhaps Quine's problem arises even for non-behaviorists. This is not the place to explore the matter.

<sup>41</sup> Roughly, the first assertion is the 'indeterminacy of translation', while the second is the 'inscrutability of reference'.

Nelson Goodman's discussion of the 'new riddle of induction' also deserves comparison with Wittgenstein's work.<sup>42</sup> Indeed, although Quine, like Wittgenstein, and unlike Goodman in his treatment of the 'new riddle', directly concerns himself with a sceptical doubt about meaning, the basic strategy of Goodman's treatment of the 'new riddle' is strikingly close to Wittgenstein's sceptical arguments. In this respect, his discussion is much closer to Wittgenstein's scepticism than is Quine's treatment of 'indeterminacy'. Although our paradigm of Wittgenstein's problem was formulated for a mathematical problem, it was emphasized that it is completely general and can be applied to any rule or word. In particular, if it were formulated for the language of color impressions, as Wittgenstein himself suggests, Goodman's 'grue' or something similar, would play the role of 'quus'.<sup>43</sup> But the problem would not be Goodman's about induction – "Why not predict that grass, which has been grue in the past, will be grue in the future?" – but Wittgenstein's about meaning: "Who is to say that in the past I didn't mean grue by 'green', so that now I should call the sky, not the grass, 'green'?" Although Goodman concentrates on the problem about induction and largely ignores the problem about meaning,<sup>44</sup> his discussions are occasionally suggestive for

<sup>42</sup> See Goodman, *Fact, Fiction, and Forecast*, p. 13b, n. 1. See also the papers in part VII ("Induction") in *Problems and Projects* (Bobbs-Merrill, Indianapolis and New York, 1972, xii + 463 pp.)

<sup>43</sup> For 'grue', see page 20 and footnotes 14 and 15 above. My memory about my own thought processes years ago is weak, but it seems likely that I may have been inspired to formulate Wittgenstein's problem in terms of 'quus' by Goodman's analogous use of 'grue'. I do remember that, at the time I first thought about the problem, I was struck by the analogy between Wittgenstein's discussion and Goodman's (as others have been as well).

<sup>44</sup> In part Goodman's discussion of the problem seems to presuppose that the extension of each predicate ('green', 'grue'), etc., is known, and that this question does not itself get entangled in the 'new riddle of induction'. Sydney Shoemaker, "On Projecting the Unprojectible," *The Philosophical Review*, vol. 84 (1975), pp. 178–219, questions whether such a separation is possible (see his concluding paragraph). I have not yet made a careful study of Shoemaker's argument.

Wittgenstein's problem as well.<sup>45</sup> In fact, I personally suspect that serious consideration of Goodman's problem, as he formulates it, may prove impossible without consideration of Wittgenstein's.<sup>46</sup>

<sup>45</sup> See his "Positionality and Pictures," *The Philosophical Review*, vol. 69 (1960), pp. 523–5, reprinted in *Problems and Projects*, pp. 402–4. See also Ullian, "More on 'Grue' and 'Gruue,'" and *Problems and Projects*, pp. 408–9 (comments on Judith Thompson).

<sup>46</sup> "Seven Structures on Similarity," *Problems and Projects*, pp. 437–46, has in places a Wittgensteinian flavor. For Goodman, as for Wittgenstein, what we call 'similar' (for Wittgenstein: even 'the same') is exhibited in our own practice and cannot explain it. (For an exposition of Wittgenstein's position, see section 3 below.)

One issue arises here. Does Wittgenstein's position depend on a denial of 'absolute similarity'? To the extent that we use 'similarity' simply to endorse the way we actually go on, it does. But it is important to see that, even if 'absolutely similar' had a fixed meaning in English, and 'similar' did not need to be filled in by a specification of the 'respects' in which things are similar, the sceptical problem would not be solved. When I learn 'plus', I could not simply give myself some finite number of examples and continue: 'Act similarly when confronted with any addition problem in the future.' Suppose that, on the ordinary meaning of 'similar' this construction is completely determinate, and that one does not hold the doctrine that various alternative ways of acting can be called 'similar', depending on how 'similar' is filled out by speaking of a respect in which one or another way of acting can be called 'similar' to what I did before. Even so, the sceptic can argue that by 'similar' I meant *quintilian*, where two actions are *quintilian* if . . . See also the discussion of 'relative identity', note 13 above.

<sup>46</sup> Briefly: Goodman insists that there is no sense that does not beg the question according to which 'grue' is 'temporal', or 'positional', and 'green' is not; if either of the pairs 'blue-green' and 'grue-bleen' is taken as primitive, the predicates of the other pair are 'temporally' definable in terms of it (see *Fact, Fiction, and Forecast*, pp. 77–80). Nevertheless, intuitively it does seem clear that 'grue' is positional in a sense that 'green' is not. Perhaps that sense can be brought out by the fact that 'green', but not 'grue', is learned (learnable?) ostensibly by a sufficient number of samples, without reference to time. It would seem that a reply to this argument should take the form: "Who is to say that it is not 'grue' that others (or even, myself in the past?) learned by such ostensive training?" But this leads directly to Wittgenstein's problem. The papers cited in the previous footnote are relevant. (It is true, however, that problems like Goodman's can arise for competing predicates that do not appear, even intuitively, to be defined positionally.)

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!"<sup>47</sup> But if I am right, a proper

<sup>47</sup> Especially for those who know some of the literature on the 'private language argument', an elaboration of this point may be useful. Much of this literature, basing itself on Wittgenstein's discussions following §243, thinks that without some external check on my identification of my own sensations, I would have no way of knowing that I have identified a given sensation correctly (in accord with my previous intentions). (The question has been interpreted to be, "How do I know I am right that this is pain?", or it might be, "How do I know that I am applying the right rule, using 'pain' as I had intended it?" See note 21 above.) But, it is argued, if I have no way of knowing (on one of these interpretations) whether I am making the right identification, it is meaningless to speak of an identification at all. To the extent that I rely on my own impressions or memories of what I meant by various sensation signs for support, I have no way of quelling these doubts. Only others, who recognize the correctness of my identification through my external behavior, can provide an appropriate external check.

A great deal could be said about the argument just obscurely summarized, which is not easy to follow even on the basis of longer presentations in the literature. But here I wish to mention one reaction: If

I really were in doubt as to whether I could identify any sensations correctly, how would a connection of my sensations with external behavior, or confirmation by others, be of any help? Surely I can identify that the relevant external behavior has taken place, or that others are confirming that I do indeed have the sensation in question, only because I can identify relevant sensory impressions (of the behavior, or of others confirming that I have identified the sensation correctly). My ability to make any identification of any external phenomenon rests on my ability to identify relevant sensory (especially visual) impressions. If I were to entertain a general doubt of my ability to identify any of my own mental states, it would be impossible to escape from it.

It is in this sense that it may appear that the argument against private language supposes that I need external help to identify my own sensations. For many presentations of the argument make it appear to depend on such a general doubt of the correctness of all my identifications of inner states. It is argued that since any identification I make needs some kind of verification for correctness, a verification of one identification of an inner state by another such identification simply raises the very same question (whether I am making a correct identification of my sensations) over again. As A. J. Ayer, in his well know exchange with Rush Rhees ("Can there be a Private Language?" *Proceedings of the Aristotelian Society*, supp. vol. 28 (1954), pp. 63-94, reprinted in Pitcher (ed.), *Wittgenstein: The Philosophical Investigations*, pp. 251-85, see especially p. 256), summarizes the argument, "His claim to recognize the object [the sensation], his belief that it really is the same, is not to be accepted unless it can be backed by further evidence. Apparently, too, this evidence must be public . . . Merely to check one private sensation by another would not be enough. For if one cannot be trusted to recognize one of them, neither can one be trusted to recognize the other." The argument concludes that I can make a genuine verification of the correctness of my identification only if I break out of the circle of 'private checks' to some publicly accessible evidence. But if I were so sceptical as to doubt all my identifications of inner states, how could anything public be of any help? Does not my recognition of anything public depend on the recognition of my inner states? As Ayer puts it (immediately following the earlier quotation), "But unless there is some thing that one is allowed to recognize, no test can ever be completed. . . I check my memory of the time at which the train is due to leave by visualizing a page of the timetable; and I am required to check this in its turn by looking up the page. [He is alluding to §265.] But unless I can trust my eyesight at this point, unless I can recognize the figures that I see written down, I am still no better off. . . Let the object to which I am attempting to refer be as public as you please. . . my assurance that I am using the word correctly . . . must in the end rest on the testimony of the senses. It is through

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?"<sup>48</sup> It is not that calling a sensation 'pain' is easy, and Wittgenstein must invent a difficulty.<sup>49</sup> 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.

hearing what other people say, or through seeing what they write, or observing their movements, that I am enabled to conclude that their use of the word agrees with mine. But if without further ado I can recognize such noises or shapes or movements, why can I not also recognize a private sensation?"

Granted that the private language argument is presented simply in this form, the objection seems cogent. Certainly it once seemed to me on some basis such as this that the argument against private language could not be right. Traditional views, which are very plausible unless they are decisively rebutted, hold that all identifications rest on the identification of sensations. The sceptical interpretation of the argument in this essay, which does not allow the notion of an identification to be taken for granted, makes the issue very different. See the discussion, on pp. 67–8 below, of an analogous objection to Hume's analysis of causation.

<sup>48</sup> So put, the problem has an obvious Kantian flavor.

<sup>49</sup> See especially the discussions of 'green' and 'grue' above, 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.

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".<sup>50</sup> Furthermore, Hume is the prime source of some ideas on the nature of mental states that Wittgenstein is most concerned to attack.<sup>51</sup> 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".<sup>52</sup> Even more forcefully, discussing the problem of the external world: "We

<sup>50</sup> 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, NJ, 1964, viii + 340 pp), p. 325.

<sup>51</sup> 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. (See the discussion of determining the meaning of 'green' with an image on p. 20 above and the corresponding discussion of the cube on pp. 42–3 above.) Of course the Wittgensteinian paradox is, among other things, a strong protest against such suppositions.

<sup>52</sup> David Hume, *A Treatise of Human Nature* (ed. L. A. Selby-Bigge,

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."<sup>53</sup> 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'.<sup>54</sup>

Clarendon Press, Oxford, 1888), Book I, Part IV, Section I (p. 183 in the Selby-Bigge edition).

<sup>53</sup> Hume, *ibid.*, Book I, Part II, 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 a 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).

<sup>54</sup> George Berkeley, *The Principles of Human Knowledge*, §§29-34. Of course

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.<sup>55</sup>

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 —

the characterization may be oversimplified, but it suffices for present purposes.

<sup>55</sup> 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 will 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.

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.<sup>56</sup> 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

<sup>56</sup> 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.



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 (p. 5), that Wittgenstein's professed inability to write a work with conventionally organized arguments and conclusions seems 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’.<sup>57</sup> 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 interlocutor (e.g. §18g; see also §195)<sup>58</sup> 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 another statement that, we must admit, is still ‘perfectly all right, properly understood’.<sup>59</sup>

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

<sup>57</sup> 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.

<sup>58</sup> §18g: “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.”

<sup>59</sup> An example of the kind of tension that can be involved appeared already above – see pp. 49–51 and note 33.

sceptic, that there is no fact as to whether I mean plus or minus. But if this is to be conceded to the sceptic, is this not the end of the matter? What can be said on behalf of our ordinary attributions of meaningful language to ourselves and to others? Has not the incredible and self-defeating conclusion, that all language is meaningless, already been drawn?

In reply we must say something about the change in Wittgenstein’s philosophy of language from the *Tractatus* to the *Investigations*. Although in detail the *Tractatus* is among the most difficult of philosophical works, its rough outlines are well known. To each sentence there corresponds a (possible) fact. If such a fact, obtains, the sentence is true; if not, false. For atomic sentences, the relation between a sentence and the fact it alleges is one of a simple correspondence or isomorphism. The sentence contains names, corresponding to objects. An atomic sentence is itself a fact, putting the names in a certain relation; and it says that (there is a corresponding fact that) the corresponding objects are in the same relation. Other sentences are (finite or infinite) truth-functions of these. Even though the details of this theory have struck some as an implausible attempt to give natural language a chimerical *a priori* structure based on logical analysis alone, similar ideas, often advanced without any specific influence from the *Tractatus*, are much alive today.<sup>60</sup>

<sup>60</sup> Donald Davidson’s influential and important theory of natural language has many features in common with the *Tractatus*, even if the underlying philosophy is different. Davidson argues that some simple, almost *a priori* considerations (not requiring detailed empirical investigation of specific natural languages) put strong constraints on the form of a theory of meaning for natural languages (it must be a finitely axiomatized Tarski-style theory of truth conditions). (Although the *form* of a theory is determined without detailed empirical investigation, for a particular language the specific theory adopted is supposed to require detailed empirical support.) The fact that a theory of meaning must have this form, it is argued, puts strong constraints on the logical form, or deep structure, of natural language – very probably that it ought to be close to classical extensional first order logic. All these ideas are close to the spirit of the *Tractatus*. In particular, like the *Tractatus*, Davidson holds (i) that truth conditions are a key element in a theory of language; (ii) that the

The simplest, most basic idea of the *Tractatus* can hardly be dismissed: a declarative sentence gets its meaning by virtue of its *truth conditions*, by virtue of its correspondence to facts that must obtain if it is true. For example, "the cat is on the mat" is understood by those speakers who realize that it is true if and only if a certain cat is on a certain mat; it is false otherwise. The presence of the cat on the mat is a fact or condition-in-the-world that would make the sentence true (express a truth) if it obtained.

So stated, the *Tractatus* picture of the meaning of declarative uncovering of a hidden deep structure of language is crucial (i) a proper theory of interpretation; (ii) that the form of the deep structure is constrained in advance by theoretical, quasi-logical considerations; (iv) that, in particular, the constraints show that the deep structure has a logical form close to that of a formal language of symbolic logic; (v) that, in particular, sentences are built up from 'atoms' by logical operators; (vi) that, in particular, the deep structure of natural language is extensional in spite of the misleading appearances of surface structure. All these ideas of the *Tractatus* are repudiated in the *Investigations*, which is hostile to any attempt to analyze language by uncovering a hidden deep structure. In this last respect, modern transformational linguistics, since Noam Chomsky, has been closer to the *Tractatus* than to the *Investigations*. (But for transformational grammarians, even the form of the theory is established by specific empirical considerations requiring detailed investigation of specific natural languages.)

See also the programs of the linguists who called themselves 'generative semanticists' and of Richard Montague. Of course many of the ideas of the *Tractatus*, or of logical atomism, have not been revived in any of these theories.

(Note: In recent transformational linguistics, 'deep structure' has a specific technical meaning. 'Generative semanticists' made the repudiation of 'deep structure' a key plank of their platform. In the preceding, it is best to take 'deep structure' in the general sense of 'underlying' structure. Anyone whose theory of language leads him to applaud the doctrine of *Tractatus* 4.002 – that the understanding of language involves countless tacit conventions, invisible to the naked eye, that disguise form – believes in deep structure in this broad sense. 'Deep structure' in the specific sense was a special theory of deep structure thus broadly defined; that is one reason why it was an appropriate term. Most recent linguistic theories that rejected 'deep structure' in the specific sense accepted it in the broader sense.)

sentences may seem not only natural but even tautological. Nonetheless, as Dummett says, "the *Investigations* contains implicitly a rejection of the classical (realist) Frege-*Tractatus* view that the general form of explanation of meaning is a statement of the truth conditions".<sup>61</sup> In the place of this view, Wittgenstein proposes an alternative rough general picture. (To call it an alternative *theory* probably goes too far. Wittgenstein disclaims (§65) any intent of offering a general account of language to rival that of the *Tractatus*. Rather we have different activities related to each other in various ways.) Wittgenstein replaces the question, "What must be the case for this sentence to be true?" by two others: first, "Under what conditions may this form of words be appropriately asserted (or denied)?"; second, given an answer to the first question, "What is the role, and the utility, in our lives of our practice of asserting (or denying) the form of words under these conditions?"

Of course Wittgenstein does not confine himself to declarative sentences, and hence to assertion and denial, as I have just done. On the contrary, any reader of the earlier parts of *Philosophical Investigations* will be aware that he is strongly concerned to deny any special primacy to assertion, or to sentences in the indicative mood. (See his early examples "Slab!", "Pillar!", etc.) This in itself plays an important role in his repudiation of the classical realist picture. Since the indicative mood is not taken as in any sense primary or basic, it becomes more plausible that the linguistic role even of utterances in the indicative mood that superficially look like assertions need not be one of 'stating facts'.<sup>62</sup> Thus, if we speak properly, we should not speak of conditions of 'assertion'.<sup>63</sup> Dummett, "Wittgenstein's Philosophy of Mathematics," p. 348 in original; reprinted in Pitcher (ed.), *Wittgenstein: The Philosophical Investigations*, pp. 446-7.

<sup>62</sup> See, for example, §304, where Wittgenstein is dealing with sensation language: "The paradox disappears only if we make a radical break with the idea that language . . . always serves the same purpose: to convey thoughts – which may be about houses, pains, good and evil, or anything else you please."

tion', but rather, more generally, of the conditions when a move (a form of linguistic expression) is to be made in the 'language game'. If, however, we allow ourselves to adopt an oversimplified terminology more appropriate to a special range of cases, we can say that Wittgenstein proposes a picture of language based, not on *truth conditions*, but on *assertability conditions* or *justification conditions*:<sup>63</sup> under what circumstances are we allowed to make a given assertion? Pictures, indeed explicit theories, of this kind are hardly unknown before

<sup>63</sup> Speaking of 'justification conditions' does not suggest the primacy of the indicative mood as much as 'assertability conditions', but it has its own drawbacks. For Wittgenstein, there is an important class of cases where a use of language properly has no independent justification other than the speaker's inclination to speak thus on that occasion (e.g. saying that one is in pain). In such cases, Wittgenstein says (§289), "To use a word without a justification (*Rechtfertigung*) does not mean to use it *zu Unrecht*." Anscombe's translation of *zu Unrecht* is not consistent. In her translation of *Philosophical Investigations*, §289, she translates it 'without right'. However, in her translation of *Remarks on the Foundations of Mathematics*, v. §33 [vii, §40], where almost exactly the same German sentence occurs, she translates it as 'wrongfully'. The German-English dictionary I have at hand (Wildhagen-Heraucourt, Brandstetter Verlag, Wiesbaden, and Allen and Unwin, London, 6th ed., 1962), translates '*zu Unrecht*' as 'unjustly, unfairly'; '*Unrecht*' in general is an 'injustice' or a 'wrong'. All this is reasonably consistent with 'wrongfully' but gives little support to 'without right', even though the idea that we have a 'right' to use a word in certain circumstances without 'justification' ('*Rechtfertigung*') is obviously in harmony with the point Wittgenstein is trying to make. However, by '*zu Unrecht*' Wittgenstein seems to mean that the use of a word without independent justification need not be a 'wrongful' use of the word — one without proper epistemic or linguistic support. On the contrary, it is essential to the workings of our language that, in some cases, such a use of language is perfectly proper. When we use the terminology of 'justification conditions', we must construe them to include such cases (where Wittgenstein would say there is no 'justification'). (Simply 'wrongly', might be a more idiomatic translation than 'wrongfully'. 'Without right' sounds to me too much as if a difficult new technical term is being introduced. The point is that '*zu Unrecht*', being a fairly ordinary German expression, should not be rendered so as to appear to be an unusual technical expression in English.) See also pp. 87-8 and note 75 below.

Wittgenstein and probably influenced him. The positivist verification theory of meaning is one of this kind. So, in a more special context, is the intuitionist account of mathematical statements. (The classical mathematician's emphasis on truth conditions is replaced by an emphasis on provability conditions.) But of course Wittgenstein's rough picture should not be identified with either of these. Its second component is distinct: granted that our language game permits a certain 'move' (assertion) under certain specifiable conditions, what is the role in our lives of such permission? Such a role must exist if this aspect of the language game is not to be idle.

Wittgenstein's alternative picture of language is already clearly suggested in the very first section of *Philosophical Investigations*. Many philosophers of mathematics — in agreement with the Augustinian conception of 'object and name' — ask such questions as, "What entities ('numbers') are denoted by numerals? What relations among these entities ('facts') correspond to numerical statements?" (Nominalistically inclined philosophers would counter, sceptically, "Can we really believe that there are such entities?") As against such a 'Platonist' conception of the problem, Wittgenstein asks that we discard any *a priori* conceptions and *look* ("Don't think, look!") at the circumstances under which numerical assertions are actually uttered, and at what roles such assertions play in our lives.<sup>64</sup> Suppose I go to the grocer with a slip marked 'five

<sup>64</sup> In some ways Frege can be taken to be the target here. It is he who insists on regarding numbers as *objects*, and on asking about the nature of these objects (even insisting that we can ask whether Julius Caesar is a number or not). On the other hand, the famous contextual principle of *Grundgesetzen der Arithmetik* (that one should ask for the signification of a sign only in the context of a sentence) and his emphasis in particular on asking how numerical expressions are actually applied is in the spirit of Wittgenstein's discussion. Perhaps the best conception of Wittgenstein's relation to Frege here is to say that Wittgenstein would regard the spirit of Frege's contextual principle as sound but would criticize Frege for using 'name of an object' as a catch-all for uses of language that are 'absolutely unlike' (§10).

red apples', and he hands over apples, reciting by heart the numerals up to five and handing over an apple as each numeral is intoned. It is under circumstances such as these that we are licensed to make utterances using numerals; the role and utility of such a license is obvious. In §§8-10, Wittgenstein imagines the letters of the alphabet, recited in alphabetical order, used in a miniature language game, just as the numbers are in this example. We have little inclination to wonder about the nature of the entities 'denoted' by the letters of the alphabet. Nevertheless, if they are used in the way described, they can properly be said to 'stand for numbers'. Indeed, to say words stand for (natural) numbers is to say that they are used as numerals, that is, used in the way described. Nevertheless the legitimacy, in its own way, of the expression 'stand for numbers' should not lead us to think of numerals as similar to expressions such as 'slab', 'pillar', and the like, except that the entities 'denoted' are not spatio-temporal. If the use of the expression 'stands for numbers' misleads in this way, it would be best to think in terms of another terminology, say, that an expression 'plays the role of a numeral'. This role, as Wittgenstein describes it, is plainly in strong contrast with the role of such expressions as 'slab', 'pillar', 'block', in the language games he describes in his early sections. (See §10.)

The case is a fine example of various aspects of Wittgenstein's technique in the *Investigations*. An important view in the philosophy of mathematics is suggested briefly almost *en passant*, almost hidden in a general discussion of the nature of language and 'language games'.<sup>65</sup> In the style discussed above,

<sup>65</sup> Paul Benacerraf, in "What Numbers Could Not Be," *The Philosophical Review*, vol. 74 (1963), pp. 47-73, see especially pp. 71-2, concludes with suggestions strikingly similar to Wittgenstein's though much of the preceding argumentation has no direct parallel in Wittgenstein. It is possible that one reason the resemblance of the views to those of a fairly well-known portion of the *Investigations* was not noticed is the *en passant* way Wittgenstein introduces the issue in the philosophy of mathematics in the context of a more general discussion. (Although I do not take it upon myself to criticize Wittgenstein in this essay, it seems to me that a great deal of further work must be done if one wishes to defend

Wittgenstein suggests that such an expression as 'stands for a number' is in order, but is dangerous if it is taken to make a certain metaphysical suggestion. In the sense this is intended by 'Platonists', one suspects him of *denying* that numerals stand for entities called 'numbers'. Most important for the present purpose, the case exemplifies the central questions he wishes to ask about the use of language. Do not look for 'entities' and 'facts' corresponding to numerical assertions, but look at the circumstances under which utterances involving numerals are made, and the utility of making them under these circumstances.

Now the replacement of truth conditions by justification conditions has a dual role in the *Investigations*. First, it offers a new approach to the problems of how language has meaning, contrasted with that of the *Tractatus*. But second, it can be applied to give an account of assertions about meaning themselves, regarded as assertions *within* our language. Recall Wittgenstein's sceptical conclusion: no facts, no truth conditions, correspond to statements such as "Jones means addition by '+'." (The present remarks about meaning and use do not in themselves provide such truth conditions. According to them, Jones now means addition by '+' if he presently intends to use the '+' sign in one way, quaddition if he intends to use it another way. But nothing is said to illuminate the question as to the nature of such an intention.)

Now if we suppose that facts, or truth conditions, are of the essence of meaningful assertion, it will follow from the sceptical conclusion that assertions that anyone ever means anything are meaningless. On the other hand, if we apply to these assertions the tests suggested in *Philosophical Investigations*, no such conclusion follows. All that is needed to legitimize assertions that someone means something is that

Wittgenstein's position here, since mathematics involves much more by way of apparently treating numbers as entities than can be covered by the simple case of counting. Perhaps some later authors can be interpreted as attempting to carry out such a project, but it is not my task to discuss these issues here.)

there be roughly specifiable circumstances under which they are legitimately assertable, and that the game of asserting them under such conditions has a role in our lives. No supposition that 'facts correspond' to those assertions is needed.

I would therefore give the following rough structure to *Philosophical Investigations* (but the breaks between parts are not sharp and to an extent are arbitrary). §§1-137 give Wittgenstein's preliminary refutation of the *Tractatus* theory of language, and suggest the rough picture he intends to put in its place. These sections come first for more than one reason. First, Wittgenstein himself once found the *Tractatus* theory natural and inevitable – Malcolm says that even in his later period he regarded it as the *only* alternative to his later work<sup>66</sup> – and sometimes he writes as if the reader will naturally be inclined to the *Tractatus* theory unless he personally intervenes to prevent it. Thus the initial sections contain a refutation, not only of the most basic and apparently inevitable theories of the *Tractatus* (such as meaning as stating facts), but also of many of its more special doctrines (such as that of a special realm of 'simples').<sup>67</sup> Wittgenstein's contrast in these initial sections between his new way of looking at matters and his old way of thinking ranges from such special views of the *Tractatus* to the nature of philosophy. This first aspect of the initial sections has, I think, been clear to most readers. Less obvious is a second aspect. The sceptical paradox is the fundamental problem of *Philosophical Investigations*. If Wittgenstein is right, we cannot begin to solve it if we remain in the grip of the natural presupposition that meaningful declarative sentences

<sup>66</sup> See Norman Malcolm, *Ludwig Wittgenstein: A Memoir*, with a biographical sketch by G. H. von Wright (Oxford University Press, London, 1958), p. 69.

<sup>67</sup> Although Wittgenstein's concern in these initial sections is primarily with his own earlier way of thinking, of course he is concerned as well with related views (the 'object and name' model of language, the picture of sentences 'as corresponding to facts', etc.) in other writers, even though these writers may have views that differ in detail from those of the *Tractatus*. He wishes to relate the discussion to larger issues as well as to his own specific views.

must purport to correspond to facts; if this is our framework, we can only conclude that sentences attributing meaning and intention are themselves meaningless. Whether or not Wittgenstein is right in thinking that the entire *Tractatus* view is a consequence of natural and apparently inevitable presuppositions, he is surely right about this fundamental part of it. The picture of correspondence-to-facts must be cleared away before we can begin with the sceptical problem.

Sections 138-242 deal with the sceptical problem and its solution. These sections – the central sections of *Philosophical Investigations* – have been the primary concern of this essay. We have not yet looked at the solution of the problem, but the astute reader will already have guessed that Wittgenstein finds a useful role in our lives for a 'language game' that licenses, under certain conditions, assertions that someone 'means such-and-such' and that his present application of a word 'accords' with what he 'meant' in the past. It turns out that this role, and these conditions, involve reference to a community. They are inapplicable to a single person considered in isolation. Thus, as we have said, Wittgenstein rejects 'private language' as early as §202.

The sections following §243 – the sections usually called 'the private language argument' – deal with the *application* of the general conclusions about language drawn in §§138-242 to the problem of sensations. The sceptical conclusion about rules, and the attendant rejection of private rules, is hard enough to swallow in general, but it seems especially unnatural in two areas. The first is mathematics, the subject of most of the preceding discussion in the present essay (and of much of Wittgenstein's in §§138-242). Do I not, in elementary mathematics, grasp rules such as that for addition, which determine all future applications? Is it not in the very nature of such rules that, once I have grasped one, I have no future choice in its application? Is not any questioning of these assertions a questioning of mathematical proof itself? And is not the grasping of a mathematical rule the solitary achievement of each mathematician independent of any interaction

with a wider community? True, others may have taught me the concept of addition, but they acted only as heuristic aids to an achievement – the 'grasping of the concept' of addition – that puts me in a special relation to the addition function. Platonists have compared the grasping of a concept to a special sense, analogous to our ordinary sensory apparatus but percipient of higher entities. But the picture does not require a special Platonic theory of mathematical objects. It depends on the observation – apparently obvious on any view – that in grasping a mathematical rule I have achieved something that depends only on my own inner state, and that is immune to Cartesian doubt about the entire external material world.<sup>68</sup>

Now another case that seems to be an obvious counterexample to Wittgenstein's conclusion is that of a sensation, or mental image. Surely I can identify these after I have felt them, and any participation in a community is irrelevant! Because these two cases, mathematics and inner experience, seem so obviously to be counterexamples to Wittgenstein's view of rules, Wittgenstein treats each in detail. The latter case is treated in the sections following §243. The former case is treated in remarks that Wittgenstein never prepared for publication, but which are excerpted in *Remarks on the Foundations of Mathematics* and elsewhere. He thinks that only if we overcome our strong inclination to ignore his general conclusions about rules can we see these two areas rightly. For this reason, the conclusions about rules are of crucial importance both to the philosophy of mathematics and to the philosophy of mind. Although in his study of sensations in

<sup>68</sup> Although Wittgenstein's views on mathematics were undoubtedly influenced by Brouwer, it is worth noting here that Brouwer's intuitionist philosophy of mathematics is, if anything, even more solipsistic than its traditional 'Platonist' rival. According to this conception, mathematics can be idealized as the isolated activity of a single mathematician ('creating subject') whose theorems are assertions about his own mental states. The fact that mathematicians form a community is irrelevant for theoretical purposes. (Indeed, Brouwer himself is said to have held mysterious 'solipsistic' views that communication is impossible. The point would remain even if we left these aside.)

§243 onward he does not simply cite his general conclusions but argues this special case afresh (he does the same for mathematics elsewhere), we will only increase our difficulties in understanding an already difficult argument if we call §243 onward 'the private language argument' and study it in isolation from the preceding material. Wittgenstein had a definite plan of organization when he placed this discussion where it is.

Of course the division is not sharp. The initial 'anti-*Treatatus*' sections contain several anticipations of the 'paradox' of §§138–242,<sup>69</sup> and even of its solution. Sections 28–36 and sections 84–9 are examples. Even the very first section of the *Investigations* can be read, with hindsight, as anticipating the problem.<sup>70</sup> Nevertheless these anticipations, being cryptic allusions to the problem in the context of the problems of earlier discussion, do not fully develop the paradox and often elide the main point into other subsidiary ones.

Consider first the anticipation in sections 84–9, especially section 86, where Wittgenstein introduces the ambiguity of rules and the possibility of an infinite regress of 'rules to interpret rules'. Knowing the central problem of *Philosophical Investigations*, it is easy to see that in these sections Wittgenstein is concerned to bring out this problem, and even to allude to part of his approach to a solution (end of §87: "The sign post is in order if, in normal circumstances, it serves its purpose"). In the context, however, Wittgenstein shades his deep paradox into a much more straightforward point – that typically

<sup>69</sup> Barry Stroud emphasized this fact to me, though the responsibility for the examples and exposition in the following paragraphs is my own.

<sup>70</sup> See: "But how does he know where and how he is to look up the word 'red' and what he is to do with the word 'five'? – Well, I assume that he acts as I described. Explanations come to an end somewhere." (§1) In hindsight, this is a statement of the basic point that I follow rules 'blindly', without any justification for the choice I make. The suggestion in the section that nothing is wrong with this situation, provided that my use of 'five', 'red', etc. fits into a proper system of activities in the community, anticipates Wittgenstein's sceptical solution, as expounded below.



uses of language do not give a precise determination of their application in all cases. (See the discussion of names in §79 – "I use the name . . . without a *fixed* meaning"; of the 'chair' (?) in §80; 'Stand roughly here' in §88.) It is true, as Wittgenstein says, that his paradox shows, among other things, that every explanation of a rule could conceivably be misunderstood, and that in this respect the most apparently precise use of language does not differ from 'rough' or 'inexact', or 'open-textured' uses. Nevertheless, surely the real point of Wittgenstein's paradox is not that the rule of addition is somehow *vague*, or leaves some cases of its application undetermined. On the contrary, the word 'plus' denotes a function whose determination is *completely* precise – in this respect it does *not* resemble the vague notions expressed by 'large', 'green', and the like. The point is the sceptical problem, outlined above, that anything in my head leaves it undetermined *what* function 'plus' (as I use it) denotes (plus or quus), what 'green' denotes (green or grue), and so on. The ordinary observation, made in abstraction from any scepticism about the meaning of 'green', that the property of greenness is itself only vaguely defined for some cases, is at best distantly related. In my opinion, Wittgenstein's sceptical arguments in no way show, in this sense, that the addition function is only vaguely defined. The addition function – as Frege would emphasize – yields one precise value for each pair of numerical arguments. This much is a theorem of arithmetic. The sceptical problem indicates no vagueness in the *concept* of addition (in the way there is vagueness in the concept of greenness), or in the word 'plus', *granting* it its usual meaning (in the way the word 'green' is vague). The sceptical point is something else.<sup>71</sup>

<sup>71</sup> Though perhaps vagueness, in the ordinary sense, enters into Wittgenstein's puzzle in this way: when a teacher *introduces* such a word as 'plus' to the learner, if he does not reduce it to more 'basic', previously learned concepts, he introduces it by a finite number of examples, plus the instructions: "Go on in the same way!" The last clause may indeed be regarded as vague, in the ordinary sense, though our grasp of the most precise concept depends on it. This type of vagueness is intimately connected with Wittgenstein's paradox.

In the sections under discussion, Wittgenstein is arguing that *any* explanation *may* fail of its purpose: if it does not in fact fail, it may work perfectly, even if the concepts involved violate the Fregean requirement of 'sharp boundaries' (§71). See §88: "If I tell someone "Stand roughly here" may not this explanation work perfectly? And cannot every other one fail too?" At least two issues are involved here: the propriety of vagueness, of violations of the Fregean requirement (actually Wittgenstein questions whether this requirement, in an absolute sense, is well-defined); and an adumbration of the sceptical paradox of the second portion (§§138–242) of the *Investigations*. In its present context, the paradox, briefly foreshadowed, is not clearly distinguished from the other considerations about vagueness and sharp boundaries. The real development of the problem is yet to come.

Similar remarks apply to the discussion of ostensive definition in §§28–36, which is part of a larger discussion of naming, one of the important topics for the first portion (§§1–137) of the *Investigations*. Wittgenstein emphasizes that ostensive definitions are always in principle capable of being misunderstood, even the ostensive definition of a color word such as 'sepia'. How someone understands the word is exhibited in the way someone goes on, "the use that he makes of the word defined". One may go on in the right way given a purely minimal explanation, while on the other hand one may go on in another way no matter how many clarifications are added, since these too can be misunderstood (a rule for interpreting a rule again; see especially §§28–9).

Much of Wittgenstein's argument is directed against the view of a special, qualitatively unique experience of understanding the ostensive definition in the right way (§§33–6). Once again Wittgenstein's real point, here in the context of naming and ostensive definition, is the sceptical paradox. The case of ostensive definition of a color ('sepia') has a special connection with the so-called 'private language argument', as developed for sensations in §§243ff. Here too, however, the argument is adumbrated so briefly, and is so much embedded



in a context of other issues, that at this stage of the argument the point can easily be lost.<sup>72</sup>

Yet another feature of the situation indicates how the ideas can be connected in a way that cuts across the indicated divisions of *Philosophical Investigations*. The first part (up to §137), as we have said, criticizes Wittgenstein's earlier picture of the nature of language and attempts to suggest another. Since Wittgenstein's sceptical solution of his paradox is possible only given his later conception of language and is ruled out by the earlier one, the discussion in the second part (§§138–242) is dependent on that of the first. The point to be made here is that, at the same time, the second part is important for an ultimate understanding of the first. Wittgen-

<sup>72</sup> In these sections, Wittgenstein does not cite examples like 'grue' or 'quus' but begins by emphasizing the ordinary possibilities for misunderstanding an ostensive definition. Many philosophers who have been influenced by Wittgenstein have happened also to be attracted to the idea that an act of ostension is ill defined unless it is accompanied by a sortal ('the entity I am pointing to' versus 'the color I am pointing to', 'the shape . . .', 'the table . . .', etc.). Then morals regarding naming and identity (as associated with 'sortal terms') are drawn from this fact. I have the impression that many of these philosophers would interpret Wittgenstein's §§28–9 as making the same point. (See, e.g., M. Dummett, *Frege* (Duckworth, London, 1973, xxv + 698 pp.), pp. 179–80, and frequently elsewhere.) However, it seems clear to me that the main point of these sections is almost the exact opposite. It should be clear from reading §29 that the idea of adding a sortal ("This number is called 'two'") is introduced by Wittgenstein's imaginary interlocutor. As against this, Wittgenstein replies that the point is in a sense correct, but that the original ostensive definition — without a sortal — is perfectly legitimate provided that it leads the learner to apply such a word as 'two' correctly in the future, while even if the sortal term is added, the possibility of future misapplication is not removed, since the sortal too may be interpreted incorrectly (and this problem cannot be removed by further explanations). Really there are two separable issues, as in the case of §§84–9. One issue is analogous to the one about vagueness in §§84–9: that an ostensive definition without an accompanying sortal is vague. The other, which clearly is the main point, is Wittgenstein's sceptical problem, presented here in terms of the possibility of misunderstanding an ostensive definition.

stein's earlier work had taken for granted a natural relation of interpretation between a thought in someone's mind and the 'fact' it 'depicts'. The relation was supposed to consist in an isomorphism between one fact (the fact that mental elements are arranged in a certain way) and another (the fact-in-the-world 'depicted'). Some of Wittgenstein's attack on this earlier idea is developed in the first part through a criticism of the notion, crucial to the *Tractatus* theory of isomorphism, of a unique decomposition of a complex into its 'ultimate' elements (see, for example, §§47–8). Clearly, however, the paradox of the second part of the *Investigations* constitutes a powerful critique of any idea that 'mental representations' uniquely correspond to 'facts', since it alleges that the components of such 'mental representations' do not have interpretations that can be 'read off' from them in a unique manner. So *a fortiori* there is no such unique interpretation of the mental 'sentences' containing them as 'depicting' one 'fact' or another.<sup>73</sup> In this way the relationship between the first and the second portions of the *Investigations* is reciprocal. In order for Wittgenstein's sceptical solution of his paradox to be intelligible, the 'realistic' or 'representational' picture of language must be undermined by another picture (in the first part). On the other hand, the paradox developed in the second part, antecedently to its solution, drives an important final nail (perhaps the crucial one) into the coffin of the representational picture.<sup>74</sup> No doubt this is one reason Wittgenstein introduces foreshadowings of the paradox already in the sections of the first part. But it also illustrates that the structural divisions I have indicated in *Philosophical Investigations* are not sharp. The investigation goes 'criss cross in every direction' (preface).

<sup>73</sup> The criticisms of the earlier ideas about 'isomorphism' are thus criticisms of a special alleged way of obtaining a unique interpretation of a mental representation. For Wittgenstein, given his earlier views, criticisms of the notion of isomorphism are thus of obvious special importance as a stage setting for his paradox. They are relatively less important as such a stage setting for someone who is not working his way out of this special milieu.

<sup>74</sup> Michael Dummett emphasized this point to me, though the responsibility for the present formulation is my own.

Wittgenstein's sceptical solution concedes to the sceptic that no 'truth conditions' or 'corresponding facts' in the world exist that make a statement like "Jones, like many of us, means addition by '+'" true. Rather we should look at how such assertions are *used*. Can this be adequate? Do we not call assertions like the one just quoted 'true' or 'false'? Can we not with propriety precede such assertions with 'It is a fact that' or 'It is not a fact that'? Wittgenstein's way with such objections is short. Like many others, Wittgenstein accepts the 'redundancy' theory of truth: to affirm that a statement is true (or presumably, to precede it with 'It is a fact that . . .') is simply to affirm the statement itself, and to say it is not true is to deny it: (' $p$ ' is true =  $p$ ). However, one might object: (a) that only utterances of certain forms are called 'true' or 'false' – questions, for example, are not – and these are so called precisely because they purport to state facts; (b) that precisely the sentences that 'state facts' can occur as components of truth-functional compounds and their meaning in such compounds is hard to explain in terms of assertability conditions alone. Wittgenstein's way with this is also short. We call something a proposition, and hence true or false, when in our language we apply the calculus of truth functions to it. That is, it is just a primitive part of our language game, not susceptible of deeper explanation, that truth functions are applied to certain sentences. For the present expository purpose it is worth noting that the sections in which he discusses the concept of truth (§§134–7) conclude the preliminary sections on the *Tractatus* and immediately precede the discussion of the sceptical paradox. They lay the final groundwork needed for that discussion.

Finally, we can turn to Wittgenstein's sceptical solution and to the consequent argument against 'private' rules. We have to see under what circumstances attributions of meaning are made and what role these attributions play in our lives. Following Wittgenstein's exhortation not to think but to look, we will not reason *a priori* about the role such statements ought to play; rather we will find out what circumstances *actually*

license such assertions and what role this license *actually* plays. It is important to realize that we are *not* looking for necessary and sufficient conditions (truth conditions) for following a rule, or an analysis of what such rule-following 'consists in'. Indeed such conditions would constitute a 'straight' solution to the sceptical problem, and have been rejected.

First, consider what is true of one person considered in isolation. The most obvious fact is one that might have escaped us after long contemplation of the sceptical paradox. It holds no terrors in our daily lives; no one actually hesitates when asked to produce an answer to an addition problem! Almost all of us unhesitatingly produce the answer '125' when asked for the sum of 68 and 57, without any thought to the theoretical possibility that a quus-like rule might have been appropriate! And we do so without justification. Of course, if asked why we said '125', most of us will say that we added 8 and 7 to get 15, that we put down 5 and carried 1 and so on. But then, what will we say if asked why we 'carried' as we do? Might our past intention not have been that 'carry' meant *quarry*; where to 'quarry' is . . . ? The entire point of the sceptical argument is that ultimately we reach a level where we act without any reason in terms of which we can justify our action. We act unhesitatingly but *blindly*.

This then is an important case of what Wittgenstein calls speaking without 'justification' ('*Rechtfertigung*'), but not 'wrongfully' ('*zu Unrecht*').<sup>75</sup> It is part of our language game of speaking of rules that a speaker may, without ultimately giving any justification, follow his own confident inclination that this way (say, responding '125') is the *right* way to

<sup>75</sup> See note 63. Note that in *Remarks on the Foundations of Mathematics*, v, §33 [VII, §40], Wittgenstein develops this point with respect to his general problem about rules, agreement, and identity, while the parallel passage in *Philosophical Investigations*, §289, is concerned with avowals of pain. This illustrates again the connection of Wittgenstein's ideas on sensation language with the general point about rules. Note also that the *RFM* passage is embedded in a context of the philosophy of mathematics: The connection of Wittgenstein's discussions of mathematics with his discussions of sensations is another theme of the present essay.

respond, rather than another way (e.g. responding '5'). That is, the 'assertability conditions' that license an individual to say that, on a given occasion, he ought to follow his rule this way rather than that, are, ultimately, that he does what he is inclined to do.

The important thing about this case is that, if we confine ourselves to looking at one person alone, his psychological states and his external behavior, this is as far as we can go. We can say that he acts confidently at each application of a rule; that he says – without further justification – that the way he acts, rather than some quus-like alternative, is *the* way to respond. There are no circumstances under which we can say that, even if he inclines to say '125', he *should* have said '5', or *vice versa*. By definition, he is licensed to give, without further justification, the answer that strikes him as natural and inevitable. Under what circumstances can he be wrong, say, following the wrong rule? No one else by looking at his mind and behavior alone can say something like, "He is wrong if he does not accord with his own past intentions"; the whole point of the sceptical argument was that there can be no facts about him in virtue of which he accords with his intentions or not. All we can say, if we consider a single person in isolation, is that our ordinary practice licenses him to apply the rule in the way it strikes him.

But of course this is *not* our usual concept of following a rule. It is by no means the case that, just because someone thinks he is following a rule, there is no room for a judgement that he is not really doing so. Someone – a child, an individual muddled by a drug – may think he is following a rule even though he is actually acting at random, in accordance with no rule at all. Alternatively, he may, under the influence of a drug, suddenly act in accordance with a quus-like rule changing from his first intentions. If there could be no justification for anyone to say of a person of the first type that his confidence that he is following some rule is misplaced, or of a person of the second type that he is no longer in accord with the rule that he previously followed, there would be little

content to our idea that a rule, or past intention, binds future choices. We are inclined to accept conditionals of such a rough type as, "If someone means addition by '+' then, if he remembers his past intention and wishes to conform to it, when he is queried about '68+57', he will answer '125'." The question is what substantive content such conditionals can have.

If our considerations so far are correct, the answer is that, if one person is considered in isolation, the notion of a rule as guiding the person who adopts it can have *no* substantive content. There are, we have seen, no truth conditions or facts in virtue of which it can be the case that he accords with his past intentions or not. As long as we regard him as following a rule 'privately', so that we pay attention to *his* justification conditions alone, all we can say is that he is licensed to follow the rule as it strikes him. This is why Wittgenstein says, "To think one is obeying a rule is not to obey a rule. 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." (§202)

The situation is very different if we widen our gaze from consideration of the rule follower alone and allow ourselves to consider him as interacting with a wider community. Others will then have justification conditions for attributing correct or incorrect rule following to the subject, and these will *not* be simply that the subject's own authority is unconditionally to be accepted. Consider the example of a small child learning addition. It is obvious that his teacher will not accept just any response from the child. On the contrary, the child must fulfill various conditions if the teacher is to ascribe to him mastery of the concept of addition. First, for small enough examples, the child must produce, almost all the time, the 'right' answer. If a child insists on the answer '7' to the query '2+3', and a '3' to '2+2', and makes various other elementary mistakes, the teacher will say to him, "You are not adding. Either you are computing another function" – I suppose he would not really talk quite this way to a child! – "or, more probably, you are as yet following no rule at all, but only giving whatever random

answer enters your head." Suppose, however, the child gets almost all 'small' addition problems right. For larger computations, the child can make more mistakes than for 'small' problems, but it must get a certain number right and, when it is wrong, it must recognizably be 'trying to follow' the proper procedure, not a gues-like procedure, even though it makes mistakes. (Remember, the teacher is not judging how accurate or *adept* the child is as an adder, but whether he can be said to be following the rule for adding.) Now, what do I mean when I say that the teacher judges that, for certain cases, the pupil must give the 'right' answer? I mean that the teacher judges that the child has given the same answer that he himself would give. Similarly, when I said that the teacher, in order to judge that the child is adding, must judge that, for a problem with larger numbers, he is applying the 'right' procedure even if he comes out with a mistaken result, I mean that he judges that the child is applying the procedure he himself is inclined to apply.

Something similar is true for adults. If someone whom I judge to have been computing a normal addition function (that is, someone whom I judge to give, when he adds, the same answer I would give), suddenly gives answers according to procedures that differ bizarrely from my own, then I will judge that something must have happened to him, and that he is no longer following the rule he previously followed. If this happens to him generally, and his responses seem to me to display little discernible pattern, I will judge him probably to have gone insane.

From this we can discern rough assertability conditions for such a sentence as "Jones means addition by 'plus'." *Jones* is entitled, subject to correction by others, provisionally to say, "I mean addition by 'plus'," whenever he has the feeling of confidence — "now I can go on!" — that he can give 'correct' responses in new cases; and *he* is entitled, again provisionally and subject to correction by others, to judge a new response to be 'correct' simply because it is the response he is inclined to give. These inclinations (both Jones's general inclination that

he has 'got it' and his particular inclination to give particular answers in particular addition problems) are to be regarded as primitive. They are not to be justified in terms of Jones's ability to interpret his own intentions or anything else. But Smith need *not* accept Jones's authority on these matters: Smith will judge Jones to mean addition by 'plus' only if he judges that Jones's answers to particular addition problems agree with those *he* is inclined to give, or, if they occasionally disagree, he can interpret Jones as at least following the proper procedure. (If Jones gives answers for very small problems disagreeing with those Smith is inclined to give, it will be difficult or impossible for Smith to interpret Jones as following the proper procedure. The same will hold if Jones's responses to larger problems are too bizarre to be errors in addition in the normal sense: for example, if he answers '5' to '68+57'.) If Jones consistently fails to give responses in agreement (in this broad sense) with Smith's, Smith will judge that he does not mean addition by 'plus'. Even if Jones did mean it in the past, the present deviation will justify Smith in judging that he has lapsed.

Sometimes Smith, by substituting some alternative interpretation for Jones's word 'plus', will be able to bring Jones's responses in line with his own. More often, he will be unable to do so and will be inclined to judge that Jones is not really following any rule at all. In all this, Smith's inclinations are regarded as just as primitive as Jones's. In no way does Smith test directly whether Jones may have in his head some rule agreeing with the one in Smith's head. Rather the point is that if, in enough concrete cases, Jones's inclinations agree with Smith's, Smith will judge that Jones is indeed following the rule for addition.

Of course if we were reduced to a babble of disagreement, with Smith and Jones asserting of each other that they are following the rule wrongly, while others disagreed with both and with each other, there would be little point to the practice just described. In fact, our actual community is (roughly) uniform in its practices with respect to addition. Any indi-

vidual who claims to have mastered the concept of addition will be judged by the community to have done so if his particular responses agree with those of the community in enough cases, especially the simple ones (and if his 'wrong' answers are not often *bizarrely* wrong, as in 's' for '68+57', but seem to agree with ours in *procedure*, even when he makes a 'computational mistake'). An individual who passes such tests is admitted into the community as an adder; an individual who passes such tests in enough other cases is admitted as a normal speaker of the language and member of the community. Those who deviate are corrected and told (usually as children) that they have not grasped the concept of addition. One who is an incorrigible deviant in enough respects simply cannot participate in the life of the community, and in communication.

Now Wittgenstein's general picture of language, as sketched above, requires for an account of a type of utterance not merely that we say under what conditions an utterance of that type can be made, but also what role and utility in our lives can be ascribed to the practice of making this type of utterance under such conditions. We say of someone else that he follows a certain rule when his responses agree with our own and deny it when they do not; but what is the utility of this practice? The utility is evident and can be brought out by considering again a man who buys something at the grocer's. The customer, when he deals with the grocer and asks for five apples, expects the grocer to count as he does, not according to some bizarre non-standard rule and so, if his dealings with the grocer involve a computation, such as '68+57', he expects the grocer's responses to agree with his own. Indeed, he may entrust the computation to the grocer. Of course the grocer may make mistakes in addition: he may even make dishonest computations. But as long as the customer attributes to him a grasp of the concept of addition, he expects that at least the grocer will not behave bizarrely, as he would if he were to follow a quus-like rule; and one can even expect that, in many cases, he will come up with the same answer the customer

would have given himself. When we pronounce that a child has mastered the rule of addition, we mean that we can entrust him to react as we do in interactions such as that just mentioned between the grocer and the customer. Our entire lives depend on countless such interactions, and on the 'game' of attributing to others the mastery of certain concepts or rules, thereby showing that we expect them to behave as we do.

This expectation is *not* infallibly fulfilled. It places a substantive restriction on the behavior of each individual, and is *not* compatible with just any behavior he may choose. (Contrast this with the case where we considered one person alone.) A deviant individual whose responses do not accord with those of the community in enough cases will not be judged, by the community, to be following its rules; he may even be judged to be a madman, following no coherent rule at all. When the community denies of someone that he is following certain rules, it excludes him from various transactions such as the one between the grocer and the customer. It indicates that it cannot rely on his behavior in such transactions.

We can restate this in terms of a device that has been common in philosophy, *inversion* of a conditional.<sup>76</sup> For example, it is important to our concept of causation that we accept some such conditional as: "If events of type *A* cause

<sup>76</sup> As will be seen immediately, inversion in this sense is a device for reversing priorities. William James summarized his famous theory of the emotions (*The Principles of Psychology*, Henry Holt & Co., New York, 1913, in 2 volumes; chapter 25 (vol. 2, 442-85), "The Emotions") by the assertion, "... the ... rational statement is that we feel sorry because we cry ... not that we cry ... because we are sorry ... " (p. 450). Many philosophies can be summed up crudely (no doubt, not really accurately) by slogans in similar form: "We do not condemn certain acts because they are immoral; they are immoral because we condemn them." "We do not accept the law of contradiction because it is a necessary truth; it is a necessary truth because we accept it (by convention)." "Fire and heat are not constantly conjoined because fire causes heat; fire causes heat because they are constantly conjoined" (Hume). "We do not all say

events of type *B*, and if an event *e* of type *A* occurs, then an event *e'* of type *B* must follow." So put, it appears that acceptance of the conditional commits us to a belief in a nexus so that, given that the causal connection between event types obtains, the occurrence of the first event *e* necessitates (by fulfilling the antecedent of the conditional), that an event *e'* of type *B* must obtain. Humeans, of course, deny the existence of such a nexus; how do they read the conditional? Essentially they concentrate on the assertability conditions of a contrapositive form of the conditional. It is not that any antecedent conditions necessitate that some event *e'* must take place; rather the conditional commits us, whenever we know that an event *e* of type *A* occurs and is not followed by an event of type *B*, to deny that there is a causal connection between the two event types. If we did make such a claim, we must now withdraw it. Although a conditional is equivalent to its contrapositive, concentration on the contrapositive reverses our priorities. Instead of seeing causal connections as primary, from which observed regularities 'flow', the Humean instead sees the regularity as primary, and – looking at the matter contrapositively – observes that we withdraw a causal hypothesis when the corresponding regularity has a definite counter-instance.

A similar inversion is used in the present instance. It is essential to our concept of a rule that we maintain some such conditional as 'If Jones means addition by '+', then if he is asked for '68+57', he will reply '125'." (Actually many clauses should be added to the antecedent to make it strictly correct, but for present purposes let us leave it in this rough form.) As in the causal case, the conditional as stated makes it appear that

12+7=19 and the like because we all grasp the concept of addition; we say we all grasp the concept of addition because we all say 12+7=19 and the like" (Wittgenstein).

The device of inversion of a conditional in the text achieves the effect of reversing priorities in a way congenial to such slogans. Speaking for myself, I am suspicious of philosophical positions of the types illustrated by the slogans, whether or not they are so crudely put.

some mental state obtains in Jones that guarantees his performance of particular additions such as '68+57' – just what the sceptical argument denies. Wittgenstein's picture of the true situation concentrates on the contrapositive, and on justification conditions. If Jones does *not* come out with '125' when asked about '68+57', we cannot assert that he means addition by '+'. Actually, of course, this is not strictly true, because our formulation of the conditional is overly loose; other conditions must be added to the antecedent to make it true. As the conditional is stated, not even the possibility of computational error is taken into account, and there are many complications not easily spelled out. The fact remains that if we ascribe to Jones the conventional concept of addition, we do not expect him to exhibit a pattern of bizarre, quus-like behavior. By such a conditional we do not mean, on the Wittgensteinian view, that any state of Jones guarantees his correct behavior. Rather by asserting such a conditional we commit ourselves, if in the future Jones behaves bizarrely enough (and on enough occasions), no longer to persist in our assertion that he is following the conventional rule of addition.

The rough conditional thus expresses a restriction on the community's game of attributing to one of its members the grasping of a certain concept: if the individual in question no longer conforms to what the community would do in these circumstances, the community can no longer attribute the concept to him. Even though, when we play this game and attribute concepts to individuals, we depict no special 'state' of their minds, we do something of importance. We take them provisionally into the community, as long as further deviant behavior does not exclude them. In practice, such deviant behavior rarely occurs.

It is, then, in such a description of the game of concept attribution that Wittgenstein's sceptical solution consists. It provides both conditions under which we are justified in attributing concepts to others and an account of the utility of this game in our lives. In terms of this account we can discuss briefly three of Wittgenstein's key concepts.



First, *agreement*. The entire 'game' we have described – that the community attributes a concept to an individual so long as he exhibits sufficient conformity, under test circumstances, to the behavior of the community – would lose its point outside a community that generally agrees in its practices. If one person, when asked to compute '68+57' answered '125', another '5', and another '13', if there was no general agreement in the community responses, the game of attributing concepts to individuals – as we have described it – could not exist. In fact of course there is considerable agreement, and deviant quus-like behavior occurs rarely. Mistakes and disagreements do occur, but these are another matter. The fact is that, extreme cases of uneducability or insanity aside, almost all of us, after sufficient training, respond with roughly the same procedures to concrete addition problems. We respond unhesitatingly to such problems as '68+57', regarding our procedure as the only comprehensible one (see, e.g., §§219, 231, 238), and we agree in the unhesitating responses we make. On Wittgenstein's conception, such agreement is essential for our game of ascribing rules and concepts to each other (see §240).

The set of responses in which we agree, and the way they interweave with our activities, is our *form of life*. Beings who agreed in consistently giving bizarre quus-like responses would share in another form of life. By definition, such another form of life would be bizarre and incomprehensible to us. ("If a lion could talk, we could not understand him" (p. 223).) However, if we can imagine the abstract possibility of another form of life (and no *a priori* argument would seem to exclude it), the members of a community sharing such a quus-like form of life could play the game of attributing rules and concepts to each other as we do. Someone would be said, in such a community, to follow a rule, as long as he agrees in his responses with the (*quus*-like) responses produced by the members of *that* community. Wittgenstein stresses the importance of agreement, and of a shared form of life, for his solution to his sceptical problem in the concluding paragraphs of the central section of *Philosophical Investigations* (§§240–2; see also the discussion of agreement on pp. 225–7).

On Wittgenstein's conception, a certain type of traditional – and overwhelmingly natural – explanation of our shared form of life is excluded. We cannot say that we all respond as we do to '68+57' because we all grasp the concept of addition in the same way, that we share common responses to particular addition problems because we share a common concept of addition. (Frege, for example, would have endorsed such an explanation, but one hardly needs to be a philosopher to find it obvious and natural.) For Wittgenstein, an 'explanation' of this kind ignores his treatment of the sceptical paradox and its solution. There is no objective fact – that we all mean addition by '+', or even that a given individual does – that explains our agreement in particular cases. Rather our license to say of each other that we mean addition by '+' is part of a 'language game' that sustains itself only because of the brute fact that we generally agree. (Nothing about 'grasping concepts' guarantees that it will not break down tomorrow.) The rough uniformities in our arithmetical behavior may or may not some day be given an explanation on the neurophysiological level, but such an explanation is not here in question.<sup>77</sup> Note again the analogy with the Humean case. Naively, we may wish to explain the observed concomitance of fire and heat by a causal, heat-producing, 'power' in the fire. The Humean alleges that any such use of causal powers to explain the regularity is meaningless. Rather we play a language game that allows us to attribute such a causal power to the fire as

<sup>77</sup> Modern transformational linguistics, inasmuch as it explains all my specific utterances by my 'grasp' of syntactic and semantic rules generating infinitely many sentences with their interpretation, seems to give an explanation of the type Wittgenstein would not permit. For the explanation is *not* in terms of my actual 'performance' as a finite (and fallible) device. It is not a purely causal (neurophysiological) explanation in the sense explained in the text, see note 22 above. On the other hand, some aspects of Chomsky's views are very congenial to Wittgenstein's conception. In particular, according to Chomsky, highly species-specific constraints – a 'form of life' – lead a child to project, on the basis of exposure to a limited corpus of sentences, a variety of new sentences for new situations. There is no *a priori* inevitability in the child's going on in the way he does, other than that this is what the species does. As was already said in note 22, the matter deserves a more extended discussion.



long as the regularity holds up. The regularity must be taken as a brute fact. So too for Wittgenstein (p. 226): "What has to be accepted, the given, is . . . forms of life."<sup>78</sup>

Finally, *criteria*. The exact interpretation and exegesis of Wittgenstein's concept of a criterion has been the subject of much discussion among students of Wittgenstein's later work. Criteria play a fundamental role in Wittgenstein's philosophy of mind: "An 'inner process' stands in need of outward criteria" (§580). Often the necessity for criteria for mental concepts has been taken, both by advocates and critics of Wittgenstein's philosophy of mind, as a fundamental *premise* of

<sup>78</sup> Can we imagine forms of life other than our own, that is, can we imagine creatures who follow rules in bizarre quus-like ways? It seems to me that there may be a certain tension in Wittgenstein's philosophy here. On the one hand, it would seem that Wittgenstein's paradox argues that there is no *a priori* reason why a creature could not follow a quus-like rule, and thus in this sense we ought to regard such creatures as conceivable. On the other hand, it is supposed to be part of our very form of life that we find it natural and, indeed, inevitable that we follow the rule for addition in the particular way that we do. (See §231: "'But surely you can see . . . ?' That is just the characteristic expression of someone who is under the compulsion of a rule.") But then it seems that we should be unable to understand 'from the inside' (cf. the notion of '*Verstehen*' in various German writers) how any creature could follow a quus-like rule. We could describe such behavior extensionally and behavioristically, but we would be unable to find it intelligible how the creature finds it natural to behave in this way. This consequence does, indeed, seem to go with Wittgenstein's conception of the matter.

Of course we can define the quus function, introduce a symbol for it, and follow the appropriate rule for computing its values. I have done so in this very essay. What it seems may be unintelligible to us is how an intelligent creature could get the very training we have for the addition function, and yet grasp the appropriate function in a quus-like way. If such a possibility were really completely intelligible to us, would we find it so inevitable to apply the plus function as we do? Yet this inevitability is an essential part of Wittgenstein's own solution to his problem.

The point is even stronger with respect to a term like 'green'. Can we grasp how someone could be presented with a number of green objects, and be told to apply the term 'green' just to 'things like these', and yet apply the term learnt as if it meant 'grue'? It would seem that if we find our own continuation to be inevitable, in some sense we cannot.

his private language argument. Critics have sometimes argued that it constitutes an undefended and indefensible verificationist assumption. Some advocates respond that if it is a verificationist premise of some sort, that form of verificationism is clearly correct.

It is not my present purpose to enter into the finer exegetical points involved in Wittgenstein's notion of a criterion,<sup>79</sup> but rather to sketch the role of the notion in the picture we have been developing. Wittgenstein's sceptical solution to his problem depends on agreement, and on checkability — on one person's ability to test whether another uses a term as he does. In our own form of life, how does this agreement come about? In the case of a term like 'table', the situation, at least in elementary cases, is simple. A child who says "table" or "That's a table" when adults see a table in the area (and does not do so otherwise) is said to have mastered the term 'table'; he says "That's a table", based on his observation, in agreement with the usage of adults, based on their observation. That is, they say, "That's a table" under like circumstances, and confirm the correctness of the child's utterances.

How does agreement emerge in the case of a term for a sensation, say 'pain'? It is not as simple as the case of 'table'. When will adults attribute to a child mastery of the avowal "I am in pain"?<sup>80</sup> The child, if he learns the avowal correctly, will utter it when he feels pain and not otherwise. By analogy with the case of 'table', it would appear that the adult should endorse this utterance if he, the adult, feels (his own? the child's?) pain. Of course we know that this is not the case. Rather the adult will endorse the child's avowal if the child's behavior (crying, agitated motion, etc.) and, perhaps, the

<sup>79</sup> One detailed attempt to enter into such issues is Rogers Albritton, "On Wittgenstein's Use of the Term 'Criterion'," in Pitcher (ed.), *Wittgenstein: The Philosophical Investigations*, pp. 231–50, reprinted with a new postscript from *The Journal of Philosophy*, vol. 56 (1959), pp. 845–57.

<sup>80</sup> Following recent (perhaps not wholly attractive) philosophical usage, I call a first person assertion that the speaker has a certain sensation (e.g. "I am in pain") an 'avowal'.

external circumstances surrounding the child, indicate that he is in pain. If a child generally avows pain under such appropriate behavioral and external circumstances and generally does not do so otherwise, the adult will say of him that he has mastered the avowal, "I am in pain."

Since, in the case of discourse on pain and other sensations, the adult's confirmation whether he agrees with the child's avowal is based on the adult's observation of the child's behavior and circumstances, the fact that such behavior and circumstances characteristic of pain exist is essential in this case to the working of Wittgenstein's sceptical solution. This, then, is what is meant by the remark, "An 'inner process' stands in need of outward criteria." Roughly speaking, outward criteria for an inner process are circumstances, observable in the behavior of an individual, which, when present, would lead others to agree with his avowals. If the individual generally makes his avowals under the right such circumstances, others will say of him that he has mastered the appropriate expression ("I am in pain," "I feel itchy," etc.). We have seen that it is part of Wittgenstein's *general* view of the workings of *all* our expressions attributing concepts that others can confirm whether a subject's responses agree with their own. The present considerations simply spell out the form this confirmation and agreement take in the case of avowals.

It should then be clear that the demand for 'outward criteria' is no verificationist or behaviorist *premise* that Wittgenstein takes for granted in his 'private language argument'. If anything, it is *deduced*, in a sense of deduction akin to Kant's.<sup>81</sup>

<sup>81</sup> See also the postscript below, note 5.

Note that it would be difficult to imagine how a causal neurophysiological explanation of the uniformities in our attributions of sensations to others (of the type mentioned on p. 97 above) could be possible if there were no 'outward' manifestations of sensations. For — except perhaps in minute or subliminal ways — the sensations of one person are causally connected to those of others only by the mediation of external signs and behavior. (I assume that 'extrasensory perception' is not in question here.) If the mediating external correlates did not exist, how could the fact

A sceptical problem is posed, and a sceptical solution to that problem is given. The solution turns on the idea that each person who claims to be following a rule can be checked by others. Others in the community can check whether the putative rule follower is or is not giving particular responses that they endorse, that agree with their own. The way they check this is, in general, a primitive part of the language game;<sup>82</sup> it need not operate the way it does in the case of 'table'.

that others agree in their judgement that a given individual has a certain sensation have a causal explanation? Causally, it would have to be a coincidence. (Similarly for the uniformities in our mathematical judgments mentioned on pp. 105–6 below.)

However, Wittgenstein does not himself seem to be particularly concerned with neurophysiological explanations of such uniformities but wants to take them as 'protophenomena' (§§654–5), where the search for an explanation is a mistake. Although I do not think such remarks are meant to rule out causal neurophysiological explanations of the uniformities, it does not appear, philosophically, that Wittgenstein wishes to *rely* on the concept of such neurophysiological explanations either.

Obviously it would be incompatible with Wittgenstein's argument to seek to 'explain' our agreement on whether a given individual is in pain in terms of our uniform 'grasp' of the concept of *pain behavior*. The fact that we agree on whether a given individual is, or is not, say, groaning, comes within the purview of Wittgenstein's sceptical arguments as much as does any other case of 'following a rule'. The causal argument sketched above is something else. (Although I have tried to avoid invoking such an argument explicitly in my discussion of 'outward criteria' in the text, since — as I said — Wittgenstein does not seem to wish to rely on such considerations, it has sometimes seemed to me that such a causal argument is implicitly involved if it is to be argued that the criteria we actually use are *essential* to our 'language game' of attributing sensations.) My discussion in this footnote and the preceding text was influenced by

a question of G. E. M. Anscombe.

<sup>82</sup> The criterion by which others judge whether a person is obeying a rule in a given instance cannot simply be his sincere inclination to say that he is, otherwise there would be no distinction between his thinking he is obeying the rule and his really obeying it (§202), and whatever he thinks is right will be right (§258). However, *after* the community judges (based on the original criteria) that he has mastered the appropriate rule, the community may (for certain rules) take the subject's sincere claim to follow it in this instance as in itself a new criterion for the correctness of

'Outward criteria' for sensations such as pain are simply the way this general requirement of our game of attributing concepts to others works out in the special case of sensations.<sup>83</sup>

his claim, without applying the original criteria. According to Wittgenstein, we do this in the case of 'I am in pain.' In the case of 'I dreamt', the terminology is originally taught to a subject who wakes up reporting certain experiences. We judge that he has mastered the rule for 'I dreamt' if he prefaces it to reports of experiences he says he had the night before. After we judge that he has mastered the language, we take 'I dreamt that such-and-such' as in itself a criterion for correctness. In both cases of 'I am in pain' and 'I dreamt', the first person utterance is new behavior that replaces the behavior that constituted the old criterion.

Reports of after-images or hallucinations are similar. We judge that someone has mastered 'I see something red' if the ordinarily utters it only when something red is present. Once we judge, however, that he has mastered this bit of language, we will accept his utterance that he sees red even when we think nothing red is present. Then we will say that he is suffering from an illusion, a hallucination, an after-image, or the like.

<sup>83</sup> One delicate point regarding sensations, and about 'criteria', ought to be noted. Wittgenstein often seems to be taken to suppose that for any type of sensation, there is an appropriate 'natural expression' of that sensation type ('pain behavior' for pain). The 'natural expression' is to be externally observable behavior 'expressing' the sensation other than, and prior to, the subject's verbal avowal that he has the sensation. If the theory of §244 that first person sensation avowals are verbal replacements for a 'primitive natural expression' of a sensation has the generality it appears to have, it would follow that Wittgenstein holds that such a 'primitive natural expression' must always exist if the first person avowal is to be meaningful. The impression is reinforced by other passages such as §§256-7. Further, the presentation of the private language argument in the present essay argues that for each rule I follow there must be a criterion - other than simply what I say - by which another will judge that I am following the rule correctly. Applied to sensations, this seems to mean that there must be some 'natural expression', or at any rate some external circumstances other than my mere inclination to say that this is the same sensation again, in virtue of which someone else can judge whether the sensation is present, and hence whether I have mastered the sensation term correctly. So the picture would be that to each statement of the form "I have sensation S" there must be an 'outward criterion' associated with S, other than the mere avowal itself, by which others recognize the presence or absence of S.

Not only professed followers of Wittgenstein but many who think of

themselves as opponents (or, at least, not followers) of Wittgenstein, seem to think that something of this kind is true. That is to say, many philosophical programs seem to suppose that all sensation types are associated with some characteristic external phenomena (behavior, causes). In this essay I have largely suppressed my own views, which are by no means always in agreement with Wittgenstein's. However, I will permit myself to remark here that any view that supposes that, in this sense, an inner process always has 'outward criteria', seems to me probably to be empirically false. It seems to me that we have sensations or sensation *qualia* that we can perfectly well identify but that have no 'natural' external manifestations; an observer cannot tell in any way whether an individual has them unless that individual avows them. Perhaps a more liberal interpretation of the private language argument - which *may* be compatible with what Wittgenstein intended - would allow that a speaker might introduce some sensation terms with no 'outward criteria' for the associated sensations beyond his own sincere avowal of them. (Hence these avowals do not 'replace' any 'natural expressions' of the sensation(s), for there are none.) There will be no way anyone else will be in any position to check such a speaker, or to agree or disagree with him. (No matter what many Wittgensteinians - or Wittgenstein - would infer here, this does not in itself entail that his avowals are regarded as infallible, nor need it in itself mean that there could not later come to be ways of checking his avowals.) However, the language of the speaker, even his language of sensations, will not have the objectionable form of a 'private language', one in which anything he calls 'right' is right. The speaker can demonstrate, for many sensations that do have 'public criteria', that he has mastered the appropriate terminology for identifying these sensations. If we agree with his responses in enough cases of various sensations, we say of him that he has mastered 'sensation language'. All this, so far, is subject to external correction. But it is a primitive part of our language game of sensations that, if an individual has satisfied criteria for a mastery of sensation language in general, we then respect his claim to have identified a new type of sensation even if the sensation is correlated with nothing publicly observable. Then the only 'public criterion' for such an avowal will be the sincere avowal itself.

How does the view sketched here liberalize the private language argument as developed in the text? In the text we argued that *for each particular rule*, if conditionals of the form "If Jones follows the rule, in this instance he will . . ." are to have any point, they must be contraposed. If the community finds that in this instance Jones is not doing . . ., he is not following the rule. Only in this 'inverted' way does the notion of my behavior as 'guided' by the rule make sense. Thus for each rule there must be an 'external check' on whether I am following it in a given instance. Perhaps §202 should be taken to assert this. But this means the

It is not my purpose here to enter in detail into the exegesis of Wittgenstein's attack on an 'object and designation' model for sensation language (§293). I am not, in fact, sure that I fully

community must have a way of telling ('criterion') whether it is being followed in a given instance, which it uses to judge the speaker's mastery of the rule. This criterion cannot be simply the speaker's own sincere inclination to follow the rule a certain way – otherwise, the conditional has no content. This condition seems to be satisfied even in those cases where, *after* the community is satisfied that the speaker has mastered the language, it lets the speaker's sincere utterance be a (or *the*) criterion for their correctness. (See note 82.) In contrast, the liberal version allows that once a speaker, judged by criteria for mastery of various rules, is accepted into the community, there should be some rules where there is no way for others to check his mastery, but where that mastery is simply presumed on the basis of his membership in the community. This is simply a primitive feature of the language game. Why should Wittgenstein not allow language games like this?

I regret that I have discussed this matter so briefly in a note. I had thought at one time to expound the 'liberal' view sketched here as the 'official' Wittgenstein doctrine, which would have facilitated an exposition at greater length in the text. Certainly it is the one Wittgenstein should have adopted in accordance with the slogan "Don't think, look!", and it really is comparable with his attack on private language. On writing the final version of this essay, however, I came to worry that passages such as §§244 and §§256–7 are highly misleading unless Wittgenstein holds something stronger.

(After writing the preceding, I found that Malcolm, in his *Thought and Knowledge* (Cornell University Press, Ithaca and London, 1977, 218 pp.), writes (p. 101), "philosophers sometimes read Wittgenstein's insistence on there being a conceptual link between statements of sensation and the primitive, natural, expressions of sensation in human behavior, as implying that there is a natural nonverbal, behavioral counterpart of every statement of sensation. Wittgenstein did not mean this, and it is obviously not true." I agree that it is not true. I think it is not true even for simple avowals invoking what we might call 'names' of sensations. ("I have sensation S.") But – what is a separate question – did Wittgenstein mean this? It seems to me that even some of Malcolm's own previous expositions of Wittgenstein have given (unintentionally?) the impression that he did, at least for simple avowals invoking 'names of sensations'. I myself have vacillated on the question. Whether or not Wittgenstein meant this, I do think that the essence of his doctrines can be captured without commitment to such a strong claim.)

understand it. But it seems likely that it relates to one aspect of our present considerations. The model of the way agreement operates with respect to a word like 'table' (perhaps a paradigm of 'object and designation') is a very simple one: the child says "Table!" when he sees that a table is present and the adult agrees if he also sees that a table is present. It is tempting to suppose that this model ought to be a general one, and that if it does not apply to the case of 'pain' we must conclude that in some sense the adult can never really confirm the correctness of the child's use of "I am in pain." Wittgenstein's suggestion is that there cannot and need not be such a demand based on generalizing the use of 'table'. No *a priori* paradigm of the way concepts ought to be applied governs all forms of life, or even our own form of life. Our game of attributing concepts to others depends on agreement. It so happens that in the case of ascribing sensation language, this agreement operates in part through 'outward criteria' for first person avowals. No further 'justification' or 'explanation' for this procedure is required; this simply is *given* as how we achieve agreement here. The important role played in our lives by the practice of attributing sensation concepts to others is evident. If I attribute mastery of the term 'pain' to someone, his sincere utterance of "I am in pain," even without other signs of pain, is sufficient to induce me to feel pity for him, attempt to aid him, and the like (or, if I am a sadist, for the opposite); and similarly in other cases.

Compare the case of mathematics. Mathematical statements are generally not about palpable entities: if they are indeed to be regarded as about 'entities', these 'entities' are generally suprasensible, eternal objects. And often mathematical statements are about the infinite. Even such an elementary mathematical truth as that any two integers have a unique sum (perhaps implicitly accepted by everyone who has mastered the concept of addition, and in any case, explicitly accepted by people with elementary sophistication as a basic property of that concept) is an assertion about infinitely many instances. All the more so is this true of the 'commutative'

law, that  $x+y=y+x$  for all  $x$  and  $y$ . Yet how does agreement operate in the case of mathematics? How do we judge of someone else that he has mastered various mathematical concepts? Our judgement, as usual, stems from the fact that he agrees with us in enough particular cases of mathematical judgements (and that, even if he disagrees, we are operating with a common procedure). We do not compare his mind with some suprasensible, infinite reality: we have seen through the sceptical paradox that this is of no help if we ask, say, whether he has mastered the concept of addition. Rather we check his observable responses to particular addition problems to see if his responses agree with ours. In more sophisticated mathematical areas, he and we accept various mathematical statements on the basis of proof; and among the conditions we require for attributing to him the mastery of our mathematical concepts is his general agreement with us on what he regards as proof. Here 'proofs' are not abstract objects laid up in a mathematical heaven (say, lengthy proofs in a formal system such as *Principia*). They are visible (or audible or palpable), concrete phenomena — marks or diagrams on paper, intelligible utterances. Proofs in this sense are not only finite objects; they are also small and clear enough to be able to judge of another man's proof whether I too would regard it as proof. That is why Wittgenstein emphasizes that proof must be *surveyable*. It must be surveyable if it is to be usable as a basis for agreement in judgements.

This parallel illuminates Wittgenstein's remark that "Finitism and behaviorism are quite similar trends. Both say, but surely, all we have here is . . . Both deny the existence of something, both with a view to escaping from a confusion." (*Remarks on the Foundations of Mathematics*, p. 63 [II, §61]) How are the two trends 'quite similar'? The finitist realizes that although mathematical statements and concepts may be about the infinite (e.g. to grasp the '+' function is to grasp an infinite table), the criteria for attributing such functions to others must be 'finite', indeed 'surveyable' — for example, we attribute mastery of the concept of addition to a child on the basis of his

agreement with us on a finite number of instances of the addition table. Similarly, though sensation language may be about 'inner' states, the behaviorist correctly affirms that attribution to others of sensation concepts rests on publicly observable (and thus on behavioral) criteria. Further, the finitist and the behaviorist are right when they deny that the relation of the infinitary mathematical or inner psychological language to its 'finite' or 'outward' criteria is an adventitious product of human frailty, one that an account of the 'essence' of mathematical or sensation language would dispense with. Mathematical finitists and psychological behaviorists, however, make parallel unnecessary moves when they deny the legitimacy of talk of infinite mathematical objects or inner states. Behaviorists either condemn talk of mental states as meaningless or illegitimate, or attempt to define it in terms of behavior. Finitists similarly regard the infinitistic part of mathematics as meaningless. Such opinions are misguided: they are attempts to repudiate our ordinary language game. In this game we are allowed, for certain purposes, to assert statements about 'inner' states or mathematical functions under certain circumstances. Although the criteria for judging that such statements are legitimately introduced are indeed behavioral (or finite), finite or behavioral statements cannot replace their role in our language as we use it.

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 *contingent* 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?<sup>84</sup> 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.<sup>85</sup> 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

<sup>84</sup> See the well-known exchange between A. J. Ayer and Rush Rhees under the title "Can there be a Private Language?" (see note 47). Both participants in the exchange assume that the 'private language argument' excludes Crusoe from language. Ayer takes this alleged fact to be fatal to Wittgenstein's argument, while Rhees takes it to be fatal to Crusoe's language. 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.

<sup>85</sup> 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?"

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."<sup>86</sup> (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

<sup>86</sup> 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.



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.<sup>87</sup>

Obviously there are countless relevant aspects of Wittgenstein's philosophy of mind that I have not discussed. About some aspects I am not clear, and others have been left untouched because of the limits of this essay.<sup>88</sup> 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

<sup>87</sup> I feel some uneasiness may remain here. Considerations of time and space, as well as the fact that I might have to abandon the role of advocate in favor of that of critic, have prevented me from carrying out a more extensive discussion of this point.

<sup>88</sup> 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* (2nd ed., Yale University Press, New Haven, 1963 xix+907 pp.), 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.

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.