SAUL A. KRIPKE

Wittgenstein on Rules and Private Language

An Elementary Exposition

Harvard University Press
Cambridge, Massachusetts

I

Introductory

Wittgenstein's celebrated argument against 'private language' has been discussed so often that the utility of yet another exposition is certainly open to question. Most of the exposition which follows occurred to the present writer some time ago, in the academic year 1962–3. At that time this approach to Wittgenstein's views struck the present writer with the force of a revelation: what had previously seemed to me to be a somewhat loose argument for a fundamentally implausible conclusion based on dubious and controversial premises now appeared to me to be a powerful argument, even if the conclusions seemed even more radical and, in a sense, more implausible, than before. I thought at that time that I had seen Wittgenstein's argument from an angle and emphasis very different from the approach which dominated standard expositions. Over the years I came to have doubts. First of all, at times I became unsure that I could formulate Wittgenstein's elusive position as a clear argument. Second, the elusive nature of the subject made it possible to interpret some of the standard literature as perhaps seeing the argument in the same way after all. More important, conversations over the years showed that, increasingly, others were seeing the argument with the emphases I preferred. Nevertheless, recent expositions by very able interpreters differ enough from the

following to make me think that a new exposition may still be of use.[1]

The common view: the argument begins at §243.

A common view of the 'private language argument' in *Philosophical Investigations* assumes that it begins with section 243, and that it continues in the sections immediately following.[2] This view takes the argument to deal primarily with a problem about 'sensation language'. Further discussion of the argument in this tradition, both in support and in criticism, emphasizes such questions as whether the argument invokes a form of the verification principle, whether the form in question is justified, whether it is applied correctly to sensation language, whether the argument rests on an exaggerated scepticism about memory, and so on. Some

[1] Looking through some of the most distinguished commentaries on Wittgenstein of the last ten or fifteen years, I find some that still treat the discussion of rules cursorily, virtually not at all, as if it were a minor topic. Others, who discuss both Wittgenstein's views on the philosophy of mathematics and his views on sensations in detail, treat the discussion of rules as if it were important for Wittgenstein's views on mathematics and logical necessity but separate it from 'the private language argument'. Since Wittgenstein has more than *one* way of arguing for a given conclusion, and even of presenting a single argument, to defend the present exegesis I need not necessarily argue that these other commentaries are in error. Indeed, they may give important and illuminating expositions of facets of the *Investigations* and its argument deemphasized or omitted in this essay. Nevertheless, in emphasis they certainly differ considerably from the present exposition.

[2] Unless otherwise specified (explicitly or contextually), references are to *Philosophical Investigations*. The small numbered units of the *Investigations* are termed 'sections' (or 'paragraphs'). Page references are used only if a section reference is not possible, as in the second part of the *Investigations*. Throughout I quote the standard printed English translation (by G. E. M. Anscombe) and make no attempt to question it except in a very few instances. *Philosophical Investigations* (x+232 pp., parallel German and English text) has undergone several editions since its first publication in 1953 but the paragraphing and pagination remain the same. The publishers are Basil Blackwell, Oxford and Macmillan, New York.

This essay does not proceed by giving detailed exegesis of Wittgenstein's text but rather develops the arguments in its own way. I recommend that the reader reread the *Investigations* in the light of the present exegesis and see whether it illuminates the text.

crucial passages in the discussion following §243 – for example, such celebrated sections as §258 and §265 – have been notoriously obscure to commentators, and it has been thought that their proper interpretation would provide the key to the 'private language argument'.

The argument's real home: before §243, with §202's conclusion.

In my view, the real 'private language argument' is to be found in the sections *preceding* §243. Indeed, in §202 *the conclusion is already stated explicitly*: "Hence it is not possible to obey a rule 'privately': otherwise thinking one was obeying a rule would be the same thing as obeying it." I do not think that Wittgenstein here thought of himself as *anticipating* an argument he was to give in greater detail later. On the contrary, the crucial considerations are all contained in the discussion leading up to the conclusion stated in §202. The sections following §243 are meant to be read in the light of the preceding discussion; difficult as they are in any case, they are much less likely to be understood if they are read in isolation.

Sensations as the special case, the ground gone over again.

The 'private language argument' as applied to *sensations* is only a special case of much more general considerations about language previously argued; sensations have a crucial role as an (apparently) convincing *counterexample* to the general considerations previously stated. Wittgenstein therefore goes over the ground again in this special case, marshalling new specific considerations appropriate to it. It should be borne in mind that *Philosophical Investigations* is not a systematic philosophical work where conclusions, once definitely established, need not be reargued. Rather the *Investigations* is written as a perpetual dialectic, where persisting worries, expressed by the voice of the imaginary interlocutor, are never definitively silenced. Since the work is not presented in the form of a deductive argument with definitive theses as conclusions, the same ground is covered repeatedly, from the point of view of various special cases and from different angles, with the hope that the entire process will help the reader see the problems rightly.

Why the Investigations reargues: a perpetual dialectic.

The basic structure of Wittgenstein's approach can be presented briefly as follows: A certain problem, or in Humean

Hume's shape: a sceptical paradox, then a sceptical solution.

terminology, a 'sceptical paradox', is presented concerning the notion of a rule. Following this, what Hume would have called a 'sceptical solution' to the problem is presented. There are two areas in which the force, both of the paradox and of its solution, are most likely to be ignored, and with respect to which Wittgenstein's basic approach is most likely to seem incredible. One such area is the notion of a mathematical rule, such as the rule for addition. The other is our talk of our own inner experience, of sensations and other inner states. In treating both these cases, we should bear in mind the basic considerations about rules and language. Although Wittgenstein has already discussed these basic considerations in considerable generality, the structure of Wittgenstein's work is such that the special cases of mathematics and psychology are not simply discussed by citing a general 'result' already established, but by going over these special cases in detail, in the light of the previous treatment of the general case. By such a discussion, it is hoped that both mathematics and the mind can be seen rightly: since the temptations to see them wrongly arise from the neglect of the same basic considerations about rules and language, the problems which arise can be expected to be analogous in the two cases. In my opinion, Wittgenstein did not view his dual interests in the philosophy of mind and the philosophy of mathematics as interests in two separate, at best loosely related, subjects, as someone might be interested both in music and in economics. Wittgenstein thinks of the two subjects as involving the same basic considerations. For this reason, he calls his investigation of the foundations of mathematics "analogous to our investigation of psychology" (p. 232). It is no accident that essentially the same basic material on rules is included in both *Philosophical Investigations* and in *Remarks on the Foundations of Mathematics*,[3] both times as

The two areas where it bites: the maths rule and inner states.

Mind and mathematics as one subject, not two.

[3] Basil Blackwell, Oxford, 1956, xix+204 pp. In the first edition of *Remarks on the Foundations of Mathematics* the editors assert (p. vi) that Wittgenstein appears originally to have intended to include some of the material on mathematics in *Philosophical Investigations*.

The third edition (1978) includes more material than earlier editions and rearranges some of the sections and divisions of earlier editions. When I wrote the present work, I used the first edition. Where the references differ, the equivalent third edition reference is given in square brackets.

the basis of the discussions of the philosophies of mind and of mathematics, respectively, which follow.

Not his own views, and why a precise statement is elusive.

In the following, I am largely trying to present Wittgenstein's argument, or, more accurately, that set of problems and arguments which I personally have gotten out of reading Wittgenstein. With few exceptions, I am *not* trying to present views of my own; neither am I trying to endorse or to criticize Wittgenstein's approach. In some cases, I have found a precise statement of the problems and conclusions to be elusive. Although one has a strong sense that there is a problem, a rigorous statement of it is difficult. I am inclined to think that Wittgenstein's later philosophical style, and the difficulty he found (see his Preface) in welding his thought into a conventional work presented with organized arguments and conclusions, is not simply a stylistic and literary preference, coupled with a *penchant* for a certain degree of obscurity,[4] but stems in part from the nature of his subject.[5]

Neither Wittgenstein's argument nor Kripke's: as it struck Kripke.

I suspect – for reasons that will become clearer later – that to attempt to present Wittgenstein's argument precisely is to some extent to falsify it. Probably many of my formulations and recastings of the argument are done in a way Wittgenstein would not himself approve.[6] So the present paper should be thought of as expounding neither 'Wittgenstein's' argument nor 'Kripke's': rather Wittgenstein's argument as it struck Kripke, as it presented a problem for him.

As I have said, I think the basic 'private language argument' *precedes* section 243, though the sections following 243 are no

[4] Personally I feel, however, that the role of stylistic considerations here cannot be denied. It is clear that purely stylistic and literary considerations meant a great deal to Wittgenstein. His own stylistic preference obviously contributes to the difficulty of his work as well as to its beauty.

[5] See the discussion of this point in pages 69–70 below.

[6] See again the same discussion in pages 69–70.

doubt of fundamental importance as well. I propose to discuss the problem of 'private language' initially without mentioning these latter sections *at all*. Since these sections are often thought to *be* the 'private language argument', to some such a procedure may seem to be a presentation of Hamlet without the prince. Even if this is so, there are many other interesting characters in the play.[7]

[7] Looking over what I have written below, I find myself worried that the reader may lose the main thread of Wittgenstein's argument in the extensive treatment of finer points. In particular, the treatment of the dispositional theory below became so extensive because I heard it urged more than once as an answer to the sceptical paradox. That discussion may contain somewhat more of Kripke's argumentation in support of Wittgenstein rather than exposition of Wittgenstein's own argument than does most of the rest of this essay. (See notes 19 and 24 for *some* of the connections. The argument is, however, inspired by Wittgenstein's original text. Probably the part with the least direct inspiration from Wittgenstein's text is the argument that our dispositions, like our actual performance, are not potentially infinite. Even this, however, obviously has its origin in Wittgenstein's parallel emphasis on the fact that we explicitly think of only finitely many cases of any rule.) The treatment below (pp. 38–39) of simplicity is an example of an objection that, as far as I know, Wittgenstein never considers himself. I think that my reply is clearly appropriate, assuming that I have understood the rest of Wittgenstein's position appropriately. I urge the reader to concentrate, on a first reading, on understanding the intuitive force of Wittgenstein's sceptical problem and to regard byways such as these as secondary.

2

The Wittgensteinian Paradox

The paradox from §201, a new form of scepticism.

In §201 Wittgenstein says, "this was our paradox: no course of action could be determined by a rule, because every course of action can be made out to accord with the rule." In this section of the present essay, in my own way I will attempt to develop the 'paradox' in question. The 'paradox' is perhaps the central problem of *Philosophical Investigations*. Even someone who disputes the conclusions regarding 'private language', and the philosophies of mind, mathematics, and logic, that Wittgenstein draws from his problem, might well regard the problem itself as an important contribution to philosophy. It may be regarded as a new form of philosophical scepticism.

Grasping the rule for addition settles indefinitely many new sums.

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

whole point of the notion that in learning to add I grasp a rule: my past intentions regarding addition determine a unique answer for indefinitely many new cases in the future.

The sum I've never done: 68+57, larger than any before.

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

Correct in two senses, arithmetical and metalinguistic.

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

The bizarre sceptic doubts only the metalinguistic sense.

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

Quus defined: plus below 57, and 5 otherwise.

which I will call 'quus' and symbolize by '⊕'. It is defined by:

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

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

The sceptic's charge: I always meant quus and now misread myself.

The sceptic claims (or feigns to claim) that I am now misinterpreting my own previous usage. By 'plus', he says, I *always meant* quus;[8] now, under the influence of some insane frenzy, or a bout of LSD, I have come to misinterpret my own previous usage.

Not logically impossible: what the mistake would consist in.

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

Wild and false, so some fact about my past usage must refute it.

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

[8] Perhaps I should make a remark about such expressions as "By 'plus' I meant quus (or plus)," "By 'green' I meant green," etc. I am not familiar with an accepted felicitous convention to indicate the object of the verb 'to mean'. There are two problems. First, if one says, "By 'the woman who discovered radium' I meant the woman who discovered radium," the object can be interpreted in two ways. It may stand for a woman (Marie Curie), in which case the assertion is true only if 'meant' is used to mean referred to (as it can be used); or it may be used to denote the *meaning* of the quoted expression, not a woman, in which case the assertion is true

Not a leap in the dark: the directions I gave myself.

Of course this bizarre hypothesis, and the references to LSD, or to an insane frenzy, are in a sense merely a dramatic device. The basic point is this. Ordinarily, I suppose that, in computing '68+57' as I do, I do not simply make an unjustified leap in the dark. I follow directions I previously gave myself that uniquely determine that in this new instance I should say '125'. What are these directions? By hypothesis, I never explicitly told myself that I should say '125' in this very instance. Nor can I say that I should simply 'do the same thing

---

with 'meant' used in the ordinary sense. Second, as is illustrated by 'referred to', 'green', 'quus', etc. above, as objects of 'meant', one must use various expressions as objects in an awkward manner contrary to normal grammar. (Frege's difficulties concerning unsaturatedness are related.) Both problems tempt one to put the object in quotation marks, like the subject; but such a usage conflicts with the convention of philosophical logic that a quotation denotes the expression quoted. Some special 'meaning marks', as proposed for example by David Kaplan, could be useful here. If one is content to ignore the first difficulty and always use 'mean' to mean denote (for most purposes of the present paper, such a reading would suit at least as well as an intensional one; often I speak as if it is a *numerical function* that is meant by plus), the second problem might lead one to nominalize the objects – 'plus' denotes the plus function, 'green' denotes greenness, etc. I contemplated using italics ("'plus' means *plus*"; "'mean' may mean *denote*"), but I decided that normally (except when italics are otherwise appropriate, especially when a neologism like 'quus' is introduced for the first time), I will write the object of 'to mean' as an ordinary roman object. The convention I have adopted reads awkwardly in the written language but sounds rather reasonable in the spoken language.

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

I always did,' if this means 'compute according to the rule exhibited by my previous examples.' That rule could just as well have been the rule for quaddition (the quus function) as for addition. The idea that in fact quaddition *is* what I meant, that in a sudden frenzy I have changed my previous usage, dramatizes the problem.

The challenge in two forms: a fact, and a justification.

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

Two conditions: the constituting fact must also justify.

The directions must be contained in the fact, or it's arbitrary.

The ground rules of our formulation of the problem should be made clear. For the sceptic to converse with me at all, we

The ground rules: present usage granted, past intentions in doubt.

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

Addition itself a genuine function, the sceptic doesn't deny it.

I put the problem in this way so as to avoid confusing questions about whether the discussion is taking place 'both inside and outside language' in some illegitimate sense.[9] If we are querying the meaning of the word 'plus', how can we use it (and variants, like 'quus') at the same time? So I suppose that the sceptic assumes that he and I agree in our *present* uses of the word 'plus': we both use it to denote addition. He does *not* – at least initially – deny or doubt that addition is a genuine function, defined on all pairs of integers, nor does he deny that we can speak of it. Rather he asks why I now believe that by 'plus' in the *past*, I meant addition rather than quaddition. If I meant the former, then to accord with my previous usage I should say '125' when asked to give the result of calculating '68 plus 57'. If I meant the latter, I should say '5'.

The present exposition tends to differ from Wittgenstein's original formulations in taking somewhat greater care to make explicit a distinction between use and mention, and between questions about present and past usage. About the present example Wittgenstein might simply ask, "How do I know that I should respond '125' to the query '68+57'?" or "How do

[9] I believe I got the phrase "both inside and outside language" from a conversation with Rogers Albritton.

Not scepticism about arithmetic: 68+57=125 may be assumed.

I know that '68+57' comes out 125?" I have found that when the problem is formulated this way, some listeners hear it as a sceptical problem about *arithmetic*: "How do I know that 68+57 is 125?" (Why not answer this question with a mathematical proof?) At least at this stage, scepticism about arithmetic should not be taken to be in question: we may assume, if we wish, that 68+57 *is* 125. Even if the question is reformulated 'metalinguistically' as "How do I know that 'plus', as I use it, denotes a function that, when applied to 68 and 57, yields 125?", one may answer, "Surely I know that 'plus' denotes the plus function and accordingly that '68 plus 57' denotes 68 plus 57. But if I know arithmetic, I know that 68 plus 57 is 125. So I know that '68 plus 57' denotes 125!" And surely, if I use language at all, I cannot doubt coherently that 'plus', as I now use it, denotes plus! Perhaps I cannot (at least at this stage) doubt this about my *present* usage. But I can doubt that my *past* usage of 'plus' denoted plus. The previous remarks – about a frenzy and LSD – should make this quite clear.

Present usage can't be doubted, past usage can.

The problem restated as a hypothesis about changed usage.

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

No past fact means no present one, but the rug stays for now.

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

we mean a certain function by 'plus' is unquestioned and unquestionable. Only *past* usages are to be questioned. Otherwise, we will be unable to *formulate* our problem.

No behaviourist limits: inner states count, and even God can't tell.

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

This feature of Wittgenstein contrasts, for example, with Quine's discussion of the 'indeterminacy of translation'.[10] There are many points of contact between Quine's discussion and Wittgenstein's. Quine, however, is more than content to assume that only behavioral evidence is to be admitted into his discussion. Wittgenstein, by contrast, undertakes an extensive introspective[11] investigation, and the results of the investiga-

[10] See W. V. Quine, *Word and Object* (MIT, The Technology Press, Cambridge, Massachusetts, 1960, xi+294 pp.), especially chapter 2, 'Translation and Meaning' (pp. 26–79). See also *Ontological Relativity and Other Essays* (Columbia University Press, New York and London, 1969, viii+165 pp.), especially the first three chapters (pp. 1–90); and see also "On the Reasons for the Indeterminacy of Translation," *The Journal of Philosophy*, vol. 67 (1970), pp. 178–83.

Quine's views are discussed further below, see pp. 55–7.

[11] I do not mean the term 'introspective' to be laden with philosophical doctrine. Of course much of the baggage that has accompanied this term would be objectionable to Wittgenstein in particular. I simply mean that he makes use, in his discussion, of our own memories and knowledge of our 'inner' experiences.

tion, as we shall see, form a key feature of his argument. Further, the way the sceptical doubt is presented is not behavioristic. It is presented from the 'inside'. Whereas Quine presents the problem about meaning in terms of a linguist, trying to guess what someone *else* means by his words on the basis of his behavior, Wittgenstein's challenge can be presented to me as a question about *myself*: was there some past fact about me – what I 'meant' by plus – that mandates what I should do now?

Posed from the inside: a question about myself, not a linguist's.

To return to the sceptic. The sceptic argues that when I answered '125' to the problem '68+57', my answer was an unjustified leap in the dark; my past mental history is equally compatible with the hypothesis that I meant quus, and therefore should have said '5'. We can put the problem this way: When asked for the answer to '68+57', I unhesitatingly and automatically produced '125', but it would seem that if previously I never performed this computation explicitly I might just as well have answered '5'. Nothing justifies a brute inclination to answer one way rather than another.

Nothing justifies a brute inclination to answer one way.

Many readers, I should suppose, have long been impatient to protest that our problem arises only because of a ridiculous model of the instruction I gave myself regarding 'addition'. Surely I did not merely give myself some finite number of examples, from which I am supposed to extrapolate the whole table ("Let '+' be the function instantiated by the following examples: . . ."). No doubt infinitely many functions are compatible with *that*. Rather I learned – and internalized instructions for – a *rule* which determines how addition is to be continued. What was the rule? Well, say, to take it in its most primitive form: suppose we wish to add x and y. Take a huge bunch of marbles. First count out x marbles in one heap. Then count out y marbles in another. Put the two heaps together and count out the number of marbles in the union thus formed. The result is $x+y$. This set of directions, I may suppose, I explicitly gave myself at some earlier time. It is engraved on my mind as on a slate. It is incompatible with the hypothesis that I meant quus. It is this set of directions, not the finite list of

The protest: I internalised a rule, and the marble recipe.

The algorithm I actually use, no different in principle.

particular additions I performed in the past, that justifies and determines my present response. This consideration is, after all, reinforced when we think what I really *do* when I add 68 and 57. I do not reply automatically with the answer '125' nor do I consult some non-existent past instructions that I should answer '125' in this case. Rather I proceed according to an *algorithm* for addition that I previously learned. The algorithm is more sophisticated and practically applicable than the primitive one just described, but there is no difference in principle.

The sceptic's reply: perhaps by 'count' I meant quount.

Despite the initial plausibility of this objection, the sceptic's response is all too obvious. True, if 'count', as I used the word in the past, referred to the act of counting (and my other past words are correctly interpreted in the standard way), then 'plus' must have stood for addition. But I applied 'count', like 'plus', to only finitely many past cases. Thus the sceptic can question my present interpretation of my past usage of 'count' as he did with 'plus'. In particular, he can claim that by 'count' I formerly meant *quount*, where to 'quount' a heap is to count it in the ordinary sense, unless the heap was formed as the union of two heaps, one of which has 57 or more items, in which case one must automatically give the answer '5'. It is clear that if in the past 'counting' meant quounting, and if I follow the rule for 'plus' that was quoted so triumphantly to the sceptic, I must admit that '68+57' must yield the answer '5'. Here I have supposed that previously 'count' was never applied to heaps formed as the union of sub-heaps either of which has 57 or more elements, but if this particular upper bound does not work, another will do. For the point is perfectly general: if 'plus' is explained in terms of 'counting', a non-standard interpretation of the latter will yield a non-standard interpretation of the former.[12]

If count meant quount, the marble rule itself gives 5.

[12] The same objection scotches a related suggestion. It might be urged that the quus function is ruled out as an interpretation of '+' because it fails to satisfy some of the laws I accept for '+' (for example, it is not associative; we could have defined it so as not even to be commutative). One might even observe that, on the natural numbers, addition is the only function that satisfies certain laws that I accept – the 'recursion equations' for +: (x) $(x+0=x)$ and (x) (y) $(x+y'=(x+y)')$ where the stroke or dash indicates successor; these equations are sometimes called a 'definition' of addition. The problem is that the other signs used in these laws (the universal quantifiers, the equality sign) have been applied in only a finite number of instances, and they can be given non-standard interpretations that will fit non-standard interpretations of '+'. Thus for example '(x)' might mean for every $x<h$, where h is some upper bound to the instances where universal instantiation has hitherto been applied, and similarly for equality.

In any event the objection is somewhat overly sophisticated. Many of us who are not mathematicians use the '+' sign perfectly well in ignorance of any explicitly formulated laws of the type cited. Yet surely we use '+' with the usual determinate meaning nonetheless. What justifies us applying the function as we do?

Another go: 'independent' becomes quindependent.

It is pointless of course to protest that I intended the result of counting a heap to be *independent* of its composition in terms of sub-heaps. Let me have said this to myself as explicitly as possible: the sceptic will smilingly reply that once again I am misinterpreting my past usage, that actually 'independent' formerly meant *quindependent*, where 'quindependent' means . . .

A rule for interpreting a rule, until I apply one blindly.

Here of course I am expounding Wittgenstein's well-known remarks about "a rule for interpreting a rule". It is tempting to answer the sceptic by appealing from one rule to another more 'basic' rule. But the sceptical move can be repeated at the more 'basic' level also. Eventually the process must stop – "justifications come to an end somewhere" – and I am left with a rule which is completely unreduced to any other. How can I justify my present application of such a rule, when a sceptic could easily interpret it so as to yield any of an indefinite number of other results? It seems that my application of it is an unjustified stab in the dark. I apply the rule *blindly*.

Normally, when we consider a mathematical rule such as addition, we think of ourselves as *guided* in our application of it to each new instance. Just this is the difference between someone who computes new values of a function and someone who calls out numbers at random. Given my past intentions regarding the symbol '+', one and only one answer

The intelligence tester's 2, 4, 6, 8 and its many continuations.

is dictated as the one appropriate to '68+57'. On the other hand, although an intelligence tester may suppose that there is only one possible continuation to the sequence 2, 4, 6, 8, . . ., mathematical and philosophical sophisticates know that an indefinite number of rules (even rules stated in terms of mathematical functions as conventional as ordinary polynomials) are compatible with any such finite initial segment. So if the tester urges me to respond, after 2, 4, 6, 8, . . ., with *the* unique appropriate next number, the proper response is that no such unique number exists, nor is there any unique (rule determined) infinite sequence that continues the given one.

Do my directions for '+' differ from the tester's?

The problem can then be put this way: Did I myself, in the directions for the future that I gave myself regarding '+', really differ from the intelligence tester? True, I may not merely stipulate that '+' is to be a function instantiated by a finite number of computations. In addition, I may give myself directions for the further computation of '+', stated in terms of other functions and rules. In turn, I may give myself directions for the further computation of these functions and rules, and so on. Eventually, however, the process must stop, with 'ultimate' functions and rules that I have stipulated for myself only by a *finite* number of examples, just as in the intelligence test. If so, is not my procedure as arbitrary as that of the man who guesses the continuation of the intelligence test? In what sense is my actual computation procedure, following an algorithm that yields '125', more justified by my past instructions than an alternative procedure that would have resulted in '5'? Am I not simply following an unjustifiable impulse?[13]

The process stops at ultimate rules fixed by finite examples.

[13] Few readers, I suppose, will by this time be tempted to appeal a determination to "go on the same way" as before. Indeed, I mention it at this point primarily to remove a possible misunderstanding of the sceptical argument, not to counter a possible reply to it. Some followers of Wittgenstein – perhaps occasionally Wittgenstein himself – have thought that his point involves a rejection of 'absolute identity' (as opposed to some kind of 'relative' identity). I do not see that this is so, whether or not doctrines of 'relative' identity are correct on other grounds. Let identity be as 'absolute' as one pleases: it holds only between each thing and itself. Then the plus function is identical with itself, and the quus function is identical with itself. None of this will tell me whether I referred to the plus function or to the quus function in the past, nor therefore will it tell me which to use in order to apply the same function now.

Wittgenstein does insist (§§215–16) that the law of identity ('everything is identical with itself') gives no way out of this problem. It should be clear enough that this is so (whether or not the maxim should be rejected as 'useless'). Wittgenstein sometimes writes (§§225–27) as if the way we give a response in a new case determines what we call the 'same', as if the meaning of 'same' varies from case to case. Whatever impression this gives, it need not relate to doctrines of relative and absolute identity. The point (which can be fully understood only after the third section of the present work) can be put this way: If someone who computed '+' as we do for small arguments gave bizarre responses, in the style of 'quus', for larger arguments, and insisted that he was 'going on the same way as before', we would not acknowledge his claim that he was 'going on in the same way' as for the small arguments. What we call the 'right' response determines what we call 'going on in the same way'. None of this in itself implies that identity is 'relative' in senses that 'relative identity' has been used elsewhere in the literature.

In fairness to Peter Geach, the leading advocate of the 'relativity' of identity, I should mention (lest the reader assume I had him in mind) that he is *not* one of those I have heard expound Wittgenstein's doctrine as dependent on a denial of 'absolute' identity.

Beyond arithmetic: the Eiffel Tower and tabair.

Of course, these problems apply throughout language and are not confined to mathematical examples, though it is with mathematical examples that they can be most smoothly brought out. I think that I have learned the term 'table' in such a way that it will apply to indefinitely many future items. So I can apply the term to a new situation, say when I enter the Eiffel Tower for the first time and see a table at the base. Can I answer a sceptic who supposes that by 'table' in the past I meant *tabair*, where a 'tabair' is anything that is a table not found at the base of the Eiffel Tower, or a chair found there? Did I think explicitly of the Eiffel Tower when I first 'grasped the concept of' a table, gave myself directions for what I meant by 'table'? And even if I did think of the Tower, cannot any directions I gave myself mentioning it be reinterpreted compatibly with the sceptic's hypothesis? Most important

The point extends to sensation words: RFM I §3.

for the 'private language' argument, the point of course applies to predicates of sensations, visual impressions, and the like, as well: "*How do I know* that in working out the series + 2 I must write "20,004, 20,006" and not "20,004, 20,008"? – (The question: "How do I know that this color is 'red'?" is similar.)" (*Remarks on the Foundations of Mathematics*, I, §3.) The passage strikingly illustrates a central thesis of this essay: that Wittgenstein regards the fundamental problems of the philosophy of mathematics and of the 'private language argument' – the problem of sensation language – as at root identical, stemming from his paradox. The whole of §3 is a succinct and beautiful statement of the Wittgensteinian paradox; indeed the whole initial section of part I of *Remarks on the Foundations of Mathematics* is a development of the problem with special reference to mathematics and logical inference. It has been supposed that all I need to do to determine my use of the word 'green' is to have an image, a sample, of green that I bring to mind whenever I apply the word in the future. When I use this to justify my application of 'green' to a new object, should not the sceptical problem be obvious to any reader of Goodman?[14] Perhaps by 'green', in the past I meant *grue*,[15] and the color image, which indeed was grue, was meant to direct me to apply the word 'green' to *grue* objects always. If the *blue* object before me now is grue, then it falls in the extension of 'green', as I meant it in the past. It is no help to suppose that in the past I stipulated that 'green' was to apply to all and only those things 'of the same color as' the sample. The sceptic can reinterpret 'same color' as same *schmolor*,[16] where things have the same schmolor if . . .

Mathematics and sensation language as one problem at root.

Green and grue: the sample in my mind was grue too.

[14] See Nelson Goodman, *Fact, Fiction, and Forecast* (3rd ed., Bobbs-Merrill, Indianapolis, 1973, xiv+131 pp.), especially ch. III, §4, pp. 72–81.

[15] The exact definition of 'grue' is unimportant. It is best to suppose that past objects were grue if and only if they were (then) green while present objects are grue if and only if they are (now) blue. Strictly speaking, this is not Goodman's original idea, but it is probably most convenient for present purposes. Sometimes Goodman writes this way as well.

[16] 'Schmolor', with a slightly different spelling, appears in Joseph Ullian, "More on 'Grue' and Grue," *The Philosophical Review*, vol. 70 (1961), pp. 386–9.

Epistemological in appearance, but no fact could settle it.

Let us return to the example of 'plus' and 'quus'. We have just summarized the problem in terms of the basis of my present particular response: what tells me that I should say '125' and not '5'? Of course the problem can be put equivalently in terms of the sceptical query regarding my present intent: nothing in my mental history establishes whether I meant plus or quus. So formulated, the problem may appear to be epistemological – how can anyone know which of these I meant? Given, however, that everything in my mental history is compatible both with the conclusion that I meant plus and with the conclusion that I meant quus, it is clear that the sceptical challenge is not really an epistemological one. It purports to show that nothing in my mental history of past behavior – not even what an omniscient God would know – could establish whether I meant plus or quus. But then it appears to follow that there was no *fact* about me that constituted my having meant plus rather than quus. How could there be, if nothing in my internal mental history or external behavior will answer the sceptic who supposes that in fact I meant quus? If there was no such thing as my meaning plus rather than quus in the past, neither can there be any such thing in the present. When we initially presented the paradox, we perforce used language, taking present meanings for granted. Now we see, as we expected, that this provisional concession was indeed fictive. There can be no fact as to what I mean by 'plus', or any other word at any time. The ladder must finally be kicked away.

No past fact, so no present one: the concession was fictive.

This, then, is the sceptical paradox. When I respond in one way rather than another to such a problem as '68+57', I can have no justification for one response rather than another. Since the sceptic who supposes that I meant quus cannot be answered, there is no fact about me that distinguishes between my meaning plus and my meaning quus. Indeed, there is no fact about me that distinguishes between my meaning a definite function by 'plus' (which determines my responses in new cases) and my meaning nothing at all.

The paradox stated: no fact, not even of meaning anything at all.

Sometimes when I have contemplated the situation, I have had something of an eerie feeling. Even now as I write, I feel

The eerie feeling: something in my mind instructs my future self.

confident that there is something in my mind – the meaning I attach to the 'plus' sign – that *instructs* me what I ought to do in all future cases. I do not *predict* what I *will* do – see the discussion immediately below – but instruct myself what I ought to do to conform to the meaning. (Were I now to make a prediction of my future behavior, it would have substantive content only because it already makes sense, in terms of the instructions I give myself, to ask whether my intentions will be conformed to or not.) But when I concentrate on what is now in my mind, what instructions can be found there? How can I be said to be acting on the basis of these instructions when I act in the future? The infinitely many cases of the table are not in my mind for my future self to consult. To say that there is a general rule in my mind that tells me how to add in the future is only to throw the problem back on to other rules that also seem to be given only in terms of finitely many cases. What can there be in my mind that I make use of when I act in the future? It seems that the entire idea of meaning vanishes into thin air.

What's in my mind to consult? Meaning vanishes into thin air.

The reply heard more than once: analyse meaning dispositionally.

Can we escape these incredible conclusions? Let me first discuss a response that I have heard more than once in conversation on this topic. According to this response, the fallacy in the argument that no fact about me constitutes my meaning plus lies in the assumption that such a fact must consist in an *occurrent* mental state. Indeed the sceptical argument shows that my entire occurrent past mental history might have been the same whether I meant plus or quus, but all this shows is that the fact that I meant plus (rather than quus) is to be analyzed *dispositionally*, rather than in terms of occurrent mental states. Since Ryle's *The Concept of Mind*, dispositional analyses have been influential; Wittgenstein's own later work is of course one of the inspirations for such analyses, and some may think that he himself wishes to suggest a dispositional solution to his paradox.

The dispositional analysis I have heard proposed is simple. To mean addition by '+' is to be disposed, when asked for any sum '$x+y$' to give the sum of x and y as the answer (in

The simple dispositional analysis: I'd have answered 125 then too.

particular, to say '125' when queried about '68+57'); to mean quus is to be disposed when queried about any arguments, to respond with their *quum* (in particular to answer '5' when queried about '68+57'). True, my actual thoughts and responses in the past do not differentiate between the plus and the quus hypotheses; but, even in the past, there were dispositional facts about me that did make such a differentiation. To say that in fact I meant plus in the past is to say – as surely was the case! – that had I been queried about '68 + 57', I *would* have answered '125'. By hypothesis I was not in fact asked, but the disposition was present none the less.

Off target: being told what I'm disposed to say doesn't justify.

To a good extent this reply immediately ought to appear to be misdirected, off target. For the sceptic created an air of puzzlement as to my *justification* for responding '125' rather than '5' to the addition problem as queried. He thinks my response is no better than a stab in the dark. Does the suggested reply advance matters? How does it *justify* my choice of '125'? What it says is: "'125' is the response you are disposed to give, and (perhaps the reply adds) it would also have been your response in the past." Well and good, I know that '125' is the response I am disposed to give (I am actually giving it!), and maybe it is helpful to be told – as a matter of brute fact – that I would have given the same response in the past. How does any of this indicate that – now *or* in the past – '125' was an answer *justified* in terms of instructions I gave myself, rather than a mere jack-in-the-box unjustified and arbitrary response? Am I supposed to justify my present belief that I meant addition, not quaddition, and hence should answer '125', in terms of a *hypothesis* about my *past* dispositions? (Do I record and investigate the past physiology of my brain?) Why am I so sure that one particular hypothesis of this kind is correct, when all my past thoughts can be construed either so that I meant plus or so that I meant quus? Alternatively, is the hypothesis to refer to my *present* dispositions alone, which would hence give the right answer by definition?

Must I justify 125 by a hypothesis about my past dispositions?

Nothing is more contrary to our ordinary view – or

The §258 test, and '' as the sign that means nothing.*

Wittgenstein's – than is the supposition that "whatever is going to seem right to me is right." (§258). On the contrary, "that only means that here we can't talk about right" (*ibid.*). A candidate for what constitutes the state of my meaning one function, rather than another, by a given function sign, ought to be such that, whatever in fact I (am disposed to) do, there is a unique thing that I *should* do. Is not the dispositional view simply an equation of performance and correctness? Assuming determinism, even if I mean to denote *no* number theoretic function in particular by the sign '*', then to the same extent as it is true for '+', it is true here that for any two arguments *m* and *n*, there is a uniquely determined answer *p* that I would give.[17] (I choose one at random, as we would normally say, but causally the answer is determined.) The difference between this case and the case of the '+' function is that in the former case, but not in the latter, my uniquely determined answer can properly be called 'right' or 'wrong'.[18]

Almost every objection boils down to this one failure.

So it does seem that a dispositional account misconceives the sceptic's problem – to find a past fact that *justifies* my present response. As a candidate for a 'fact' that determines what I mean, it fails to satisfy the basic condition on such a candidate, stressed above on p. 11, that it should *tell* me what I ought to do in each new instance. Ultimately, almost all objections to the dispositional account boil down to this one. However, since the dispositionalist does offer a popular

[17] We will see immediately below that for arbitrarily large *m* and *n*, this assertion is not really true even for '+'. That is why I say that the assertion is true for '+' and the meaningless '*' 'to the same extent'.

[18] I might have introduced '*' to mean nothing in particular even though the answer I arbitrarily choose for '*m***n*' is, through some quirk in my brain structure, uniquely determined independently of the time and other circumstances when I am asked the question. It might, in addition, even be the case that I consciously resolve, once I have chosen a particular answer to '*m***n*', to stick to it if the query is repeated for any particular case, yet nevertheless I think of '*' as meaning no function in particular. What I will not say is that my particular answer is 'right' or 'wrong' in terms of the *meaning* I assigned to '*', as I will for '+', since there is no such meaning.

candidate for what the fact as to what I mean might be, it is worth examining some problems with the view in more detail.

As I said, probably some have read Wittgenstein himself as favoring a dispositional analysis. I think that on the contrary, although Wittgenstein's views have dispositional elements, any such analysis is inconsistent with Wittgenstein's view.[19]

[19] Russell's *The Analysis of Mind* (George Allen and Unwin, London, in the Muirhead Library of Philosophy, 310 pp.) already gives dispositional analyses of certain mental concepts: see especially, Lecture III, "Desire and Feeling," pp. 58–76. (The object of a desire, for example, is roughly defined as that thing which, when obtained, will cause the activity of the subject due to the desire to cease.) The book is explicitly influenced by Watsonian behaviorism; see the preface and the first chapter. I am inclined to conjecture that Wittgenstein's philosophical development was influenced considerably by this work, both in the respects in which he sympathizes with behavioristic and dispositional views, and to the extent that he opposes them. I take *Philosophical Remarks* (Basil Blackwell, Oxford, 1975, 357 pp., translated by R. Hargreaves and R. White), §§21ff., to express a rejection of Russell's theory of desire, as stated in Lecture III of *The Analysis of Mind*. The discussion of Russell's theory played, I think, an important role in Wittgenstein's development: the problem of the relation of a desire, expectation, etc., to its object ('intentionality') is one of the important forms Wittgenstein's problem about meaning and rules takes in the *Investigations*. Clearly the sceptic, by proposing his bizarre interpretations of what I previously meant, can get bizarre results as to what (in the present) does, or does not, satisfy my past desires or expectations, or what constitutes obedience to an order I gave. Russell's theory parallels the dispositional theory of meaning in the text by giving a causal dispositional account of desire. Just as the dispositional theory holds that the value I meant '+' to have for two particular arguments *m* and *n* is, by definition, the answer I would give if queried about '*m*+*n*', so Russell characterizes the thing I desired as the thing which, were I to get it, would quiet my 'searching' activity. I think that even in the *Investigations*, as in *Philosophical Remarks* (which stems from an earlier period), Wittgenstein still rejects Russell's dispositional theory because it makes the relation between a desire and its object an 'external' relation (*PR*, §21), although in the *Investigations*, unlike *Philosophical Remarks*, he no longer bases this view on the 'picture theory' of the *Tractatus*. Wittgenstein's view that the relation between the desire (expectation, etc.) and its object must be 'internal', not 'external',

The criterion: read the function off my dispositions.

First, we must state the simple dispositional analysis. It gives a criterion that will tell me what number theoretic function φ I mean by a binary function symbol 'f', namely: The referent φ of 'f' is that unique binary function φ such that I am disposed, if queried about '$f(m, n)$', where 'm' and 'n' are numerals denoting particular numbers m and n, to reply 'p', where 'p' is a numeral denoting $\varphi(m, n)$. The criterion is meant to enable us to 'read off' which function I mean by a given function symbol from my disposition. The cases of addition and quaddition above would simply be special cases of such a scheme of definition.[20]

First difficulty: my dispositions are finite too.

The dispositional theory attempts to avoid the problem of the finiteness of my actual past performance by appealing to a disposition. But in doing so, it ignores an obvious fact: not only my actual performance, but also the totality of my dispositions, is finite. It is not true, for example, that if queried about the sum of any two numbers, no matter how large, I will reply with their actual sum, for some pairs of numbers are

---

parallels corresponding morals drawn about meaning in my text below (the relation of meaning and intention to future action is 'normative, not descriptive', p. 37 below). Sections 429–65 discuss the fundamental problem of the *Investigations* in the form of 'intentionality'. I am inclined to take §440 and §460 to refer obliquely to Russell's theory and to reject it.

Wittgenstein's remarks on machines (see pp. 33–4 and note 24 below) also express an explicit rejection of dispositional and causal accounts of meaning and following a rule.

[20] Actually such a crude definition is quite obviously inapplicable to functions that I can define but cannot compute by any algorithm. Granted Church's thesis, such functions abound. (See the remark on Turing machines in footnote 24 below.) However, Wittgenstein himself does not consider such functions when he develops his paradox. For symbols denoting such functions the question "What function do I mean by the symbol?" makes sense; but the usual Wittgensteinian paradox (any response, not just the one I give, accords with the rule) makes no sense, since there need be no response that I give if I have no procedure for computing values of the function. Nor does a dispositional account of what I mean make sense. – This is not the place to go into such matters: for Wittgenstein, it may be connected with his relations to finitism and intuitionism.

Quaddition redefined to agree wherever I can add at all.

simply too large for my mind – or my brain – to grasp. When given such sums, I may shrug my shoulders for lack of comprehension; I may even, if the numbers involved are large enough, die of old age before the questioner completes his question. Let 'quaddition' be redefined so as to be a function which agrees with addition for all pairs of numbers small enough for me to have any disposition to add them, and let it diverge from addition thereafter (say, it is 5). Then, just as the sceptic previously proposed the hypothesis that I meant quaddition in the old sense, now he proposes the hypothesis that I meant quaddition in the new sense. A dispositional account will be impotent to refute him. As before, there are infinitely many candidates the sceptic can propose for the role of quaddition.

Idealise: ceteris paribus I'd give the sum.

I have heard it suggested that the trouble arises solely from too crude a notion of disposition: *ceteris paribus*, I surely will respond with the sum of any two numbers when queried. And *ceteris paribus* notions of dispositions, not crude and literal notions, are the ones standardly used in philosophy and in science. Perhaps, but how should we flesh out the *ceteris paribus* clause? Perhaps as something like: if my brain had been stuffed with sufficient extra matter to grasp large enough numbers, and if it were given enough capacity to perform such a large addition, and if my life (in a healthy state) were prolonged enough, then given an addition problem involving two large numbers, m and n, I would respond with their sum, and not with the result according to some quus-like rule. But how can we have any confidence of this? How in the world can I tell what would happen if my brain were stuffed with extra brain matter, or if my life were prolonged by some magic elixir? Surely such speculation should be left to science fiction writers and futurologists. We have no idea what the results of such experiments would be. They might lead me to go insane, even to behave according to a quus-like rule. The outcome really is obviously indeterminate, failing further specification of these magic mind-expanding processes; and even with such specifications, it is highly speculative. But of course what the

Fleshing out the clause: extra brain matter and a longer life.

Its real meaning: if I carried out my intentions, I'd give the sum.

ceteris paribus clause really means is something like this: If I somehow were to be given the means to carry out my intentions with respect to numbers that presently are too long for me to add (or to grasp), and if I were to carry out these intentions, then if queried about '$m+n$' for some big m and n, I would respond with their sum (and not with their quum). Such a counterfactual conditional is true enough, but it is of no help against the sceptic. It presupposes a prior notion of my having an intention to mean one function rather than another by '+'. It is in virtue of a fact of this kind about me that the conditional is true. But of course the sceptic is challenging the existence of just such a fact; his challenge must be met by specifying its nature. Granted that I mean addition by '+', then of course if I were to act in accordance with my intentions, I would respond, given any pair of numbers to be combined by '+', with their sum; but equally, granted that I mean quaddition, if I were to act in accordance with my intentions, I would respond with the quum. One cannot favor one conditional rather than another without circularity.

Granted I meant quaddition, I'd give the quum: circular.

The recap: idealised dispositions presuppose the function meant.

Recapitulating briefly: if the dispositionalist attempts to define which function I meant as the function determined by the answer I am disposed to give for arbitrarily large arguments, he ignores the fact that my dispositions extend to only finitely many cases. If he tries to appeal to my responses under idealized conditions that overcome this finiteness, he will succeed only if the idealization includes a specification that I will still respond, under these idealized conditions, according to the infinite table of the function I actually meant. But then the circularity of the procedure is evident. The idealized dispositions are determinate only because it is already settled which function I meant.

Second difficulty: dispositions to make mistakes.

The dispositionalist labors under yet another, equally potent, difficulty, which was foreshadowed above when I recalled Wittgenstein's remark that, if 'right' makes sense, it cannot be the case that whatever seems right to me is (by definition) right. Most of us have dispositions to make

Forgetting to carry: a mistake, or a different function?

mistakes.[21] For example, when asked to add certain numbers some people forget to 'carry'. They are thus disposed, for these numbers, to give an answer differing from the usual addition table. Normally, we say that such people have made a *mistake*. That means, that for them as for us, '+' means addition, but for certain numbers they are not disposed to give the answer they *should* give, if they are to accord with the table of the function they actually *meant*. But the dispositionalist cannot say this. According to him, the function someone means is to be *read off* from his dispositions; it cannot be

[21] However, in the slogan quoted and in §202, Wittgenstein seems to be more concerned with the question, "Am I right in thinking that I am still applying the same rule?" than with the question "Is my application of the rule right?" Relatively few of us have the disposition – as far as I know – bizarrely to cease to apply a given rule if once we were applying it. Perhaps there is a corrosive substance present in my brain already (whose action will be 'triggered' if I am given a certain addition problem) that will lead me to forget how to add. I might, once this substance is secreted, start giving bizarre answers to addition problems – answers that conform to a quus-like rule, or to no discernible pattern at all. Even if I do think that I am following the same rule, in fact I am not.

Now, when I assert that I definitely mean addition by 'plus', am I making a *prediction* about my future behavior, asserting that there is no such corrosive acid? To put the matter differently: I assert that the present meaning I give to '+' determines values for arbitrarily large amounts. I do *not* predict that I will come out with these values, or even that I will use anything like the 'right' procedures to get them. A disposition to go berserk, to change the rule, etc., may be in me already, waiting to be triggered by the right stimulus. I make no assertion about such possibilities when I say that my use of the '+' sign determines values for every pair of arguments. Much less do I assert that the values I will come out with under these circumstances are, by definition, the values that accord with what is meant.

These possibilities, and the case mentioned above with '*', when I am disposed to respond even though I follow no rule from the beginning, should be borne in mind in addition to the garden-variety possibility of error mentioned in the text. Note that in the case of '*', it seems intuitively possible that I could be under the impression that I was following a rule even though I was following none – see the analogous case of reading on pp. 45–6 below, in reference to §166.

Skaddition: the flawless skadder where we see a fallible adder.

presupposed in advance which function is meant. In the present instance a certain unique function (call it 'skaddition') corresponds in its table exactly to the subject's dispositions, including his dispositions to make mistakes. (Waive the difficulty that the subject's dispositions are finite: suppose he has a disposition to respond to any pair of arguments.) So, where common sense holds that the subject means the same addition function as everyone else but systematically makes computational mistakes, the dispositionalist seems forced to hold that the subject makes no computational mistakes, but means a non-standard function ('skaddition') by '+'. Recall that the dispositionalist held that we would detect someone who meant quus by '+' *via* his disposition to respond with '5' for arguments $\geqslant 57$. In the same way, he will 'detect' that a quite ordinary, though fallible, subject means some non-standard function by '+'.

No clause screens out the mistakes without presupposing the function.

Once again, the difficulty cannot be surmounted by a *ceteris paribus* clause, by a clause excluding 'noise', or by a distinction between 'competence' and 'performance'. No doubt a disposition to give the true sum in response to each addition problem is part of my 'competence', if by this we mean simply that such an answer accords with the rule I intended, or if we mean that, if all my dispositions to make mistakes were removed, I would give the correct answer. (Again I waive the finiteness of my capacity.) But a disposition to make a mistake is simply a disposition to *give an answer other than the one that accords with the function I meant*. To presuppose this concept in the present discussion is of course viciously circular. If I meant addition, my 'erroneous' actual disposition is to be ignored; if I meant skaddition, it should not be. Nothing in the notion of my 'competence' as thus defined can possibly tell me which alternative to adopt.[22] Alternatively, we might try to specify

[22] Lest I be misunderstood, I hope it is clear that in saying this I do not myself reject Chomsky's competence–performance distinction. On the contrary, I personally find that the familiar arguments for the distinction (and for the attendant notion of grammatical rule) have great persuasive force. The present work is intended to expound my understanding of Wittgenstein's position, not my own; but I certainly do not mean, exegetically, to assert that Wittgenstein himself would reject the distinction. But what *is* important here is that the notion of 'competence' is itself not a dispositional notion. It is normative, not descriptive, in the sense explained in the text.

The point is that our understanding of the notion of 'competence' is dependent on our understanding of the idea of 'following a rule', as is argued in the discussion above. Wittgenstein would reject the idea that 'competence' can be defined in terms of an idealized dispositional or mechanical model, and used without circularity to explicate the notion of following a rule. Only after the sceptical problem about rules has been resolved can we *then* define 'competence' in terms of rule-following. Although notions of 'competence' and 'performance' differ (at least) from writer to writer, I see no reason why linguists need assume that 'competence' is defined prior to rule-following. Although the remarks in the text warn against the use of the 'competence' notion as a solution to our problem, in no way are they arguments against the notion itself.

Nevertheless, given the sceptical nature of Wittgenstein's solution to his problem (as this solution is explained below), it is clear that if Wittgenstein's standpoint is accepted, the notion of 'competence' will be seen in a light radically different from the way it implicitly is seen in much of the literature of linguistics. For *if* statements attributing rule-following are neither to be regarded as stating facts, nor to be thought of as *explaining* our behavior (see section 3 below), it would seem that the *use* of the ideas of rules and of competence in linguistics needs serious reconsideration, even if these notions are not rendered 'meaningless'. (Depending on one's standpoint, one might view the tension revealed here between modern linguistics and Wittgenstein's sceptical critique as casting doubt on the linguistics, or on Wittgenstein's sceptical critique – or both.) These questions would arise even if, as throughout the present text, we deal with rules, like addition, that are stated explicitly. These rules we think of ourselves as grasping consciously; in the absence of Wittgenstein's sceptical arguments, we would see no problem in the assumption that each particular answer we produce is justified by our 'grasp' of the rules. The problems are compounded if, as in linguistics, the rules are thought of as tacit, to be reconstructed by the scientist and *inferred* as an *explanation* of behavior. The matter deserves an extended discussion elsewhere. (See also pp. 97 to 99 and n. 77 below.)

the 'noise' to be ignored without presupposing a prior notion of which function is meant. A little experimentation will reveal the futility of such an effort. Recall that the subject has a

Specifying the noise: correction by others, and why it's circular.

systematic disposition to forget to carry in certain circumstances: he tends to give a uniformly erroneous answer when well rested, in a pleasant environment free of clutter, etc. One cannot repair matters by urging that the subject would eventually respond with the right answer after correction by others. First, there are uneducable subjects who will persist in their error even after persistent correction. Second, what is meant by 'correction by others'? If it means rejection by others of 'wrong' answers (answers that do not accord with the rule the speaker means) and suggestion of the right answer (the answer that does accord), then again the account is circular. If random intervention is allowed (that is, the 'corrections' may be arbitrary, whether they are 'right' or 'wrong'), then, although educable subjects may be induced to correct their wrong answers, suggestible subjects may also be induced to replace their correct answers with erroneous ones. The amended dispositional statement will, then, provide no criterion for the function that is really meant.

The upshot: dispositions don't determine the function meant.

The dispositional theory, as stated, assumes that which function I meant is determined by my dispositions to compute its values in particular cases. In fact, this is not so. Since dispositions cover only a finite segment of the total function and since they may deviate from its true values, two individuals may agree on their computations in particular cases even though they are actually computing different functions. Hence the dispositional view is not correct.

The machine variant: Dummett's question about freedom of choice.

In discussions, I have sometimes heard a variant of the dispositional account. The argument goes as follows: the sceptic argues, in essence, that I am free to give any new answer to an addition problem, since I can always interpret my previous intentions appropriately. But how can this be? As Dummett put the objection: "A machine can follow this rule; whence does a human being gain a freedom of choice in this matter which a machine does not possess?"[23] The objection is

[23] M. A. E. Dummett, "Wittgenstein's Philosophy of Mathematics," *The Philosophical Review*, vol. 68 (1959), pp. 324–48, see p. 331; reprinted in George Pitcher (ed.), *Wittgenstein: The Philosophical Investigations* (Macmillan, 1966, pp. 420–47), see p. 428. The quoted objection need not necessarily be taken to express Dummett's own ultimate view of the matter.

really a form of the dispositional account, for that account can be viewed as if it interpreted us as machines, whose output mechanically yields the correct result.

Build the machine, and its answer is the one I intended.

We can interpret the objector as arguing that the rule can be *embodied* in a machine that computes the relevant function. If I build such a machine, it will simply grind out the right answer, in any particular case, to any particular addition problem. The answer that the machine would give is, then, the answer that I intended.

'Machine' is ambiguous: first, the program I draw up.

The term 'machine' is here, as often elsewhere in philosophy, ambiguous. Few of us are in a position to build a machine or draw up a program to embody our intentions; and if a technician performs the task for me, the sceptic can ask legitimately whether the technician has performed his task correctly. Suppose, however, that I am fortunate enough to be such an expert that I have the technical facility required to embody my own intentions in a computing machine, and I state that the machine is *definitive* of my own intentions. Now the word 'machine' here may refer to any one of various things. It may refer to a machine *program* that I draw up, embodying my intentions as to the operation of the machine. Then exactly the same problems arise for the program as for the original symbol '+': the sceptic can feign to believe that the program, too, ought to be interpreted in a quus-like manner.

Calling the program abstract only moves the question.

To say that a program is not something that I wrote down on paper, but an abstract mathematical object, gets us no further. The problem then simply takes the form of the question: what program (in the sense of abstract mathematical object) corresponds to the 'program' I have written on paper (in accordance with the way I meant it)? ('Machine' often seems to mean a program in one of these senses: a Turing 'machine', for example, would be better called a 'Turing program'.) Finally, however, I may build a concrete machine, made of metal and

The physical machine still needs a coding, and that's reinterpretable.

gears (or transistors and wires), and declare that it embodies the function I intend by '+': the values that it gives are the values of the function I intend. However, there are several problems with this. First, even if I say that the machine embodies the function in this sense, I must do so in terms of instructions (machine 'language', coding devices) that tell me how to interpret the machine; further, I must declare explicitly that the function always takes values as given, in accordance with the chosen code, by the machine. But then the sceptic is free to interpret all these instructions in a non-standard, 'quus-like' way. Waiving this problem, there are two others – here is where the previous discussion of the dispositional view comes in. I cannot really insist that the values of the function are given by the machine. First, the machine is a finite object, accepting only finitely many numbers as input and yielding only finitely many as output – others are simply too big. Indefinitely many programs extend the actual finite behavior of the machine. Usually this is ignored because the designer of the machine intended it to fulfill just one program, but in the present context such an approach to the intentions of the designer simply gives the sceptic his wedge to interpret in a non-standard way. (Indeed, the appeal to the designer's program makes the physical machine superfluous; only the program is really relevant. The machine as physical object is of value only if the intended function can somehow be read off from the physical object alone.) Second, in practice it hardly is likely that I really intend to entrust the values of a function to the operation of a physical machine, even for that finite portion of the function for which the machine can operate. Actual machines can *malfunction*: through melting wires or slipping gears they may give the wrong answer. How is it determined when a malfunction occurs? By reference to the program of the machine, as intended by its designer, not simply by reference to the machine itself. Depending on the intent of the designer, any particular phenomenon may or may not count as a machine 'malfunction'. A programmer with suitable intentions might even have intended to make use

Waiving that, finitude: many programs extend its behaviour.

And machines malfunction: what counts as one?

Malfunction is fixed by the program, so the program is the target.

of the fact that wires melt or gears slip, so that a machine that is 'malfunctioning' for me is behaving perfectly for him. Whether a machine ever malfunctions and, if so, when, is not a property of the machine itself as a physical object but is well defined only in terms of its program, as stipulated by its designer. Given the program, once again the physical object is superfluous for the purpose of determining what function is meant. Then, as before, the sceptic can concentrate his objections on the program. The last two criticisms of the use of the physical machine as a way out of scepticism – its finitude and the possibility of malfunction – obviously parallel two corresponding objections to the dispositional account.[24]

[24] Wittgenstein discusses machines explicitly in §§193–5. See the parallel discussion in *Remarks on the Foundations of Mathematics*, part I, §§118–30, especially §§119–26; see also, e.g., II [III], §87, and III [IV], §§48–9 there. The criticisms in the text of the dispositional analysis and of the use of machines to solve the problem are inspired by these sections. In particular, Wittgenstein himself draws the distinction between the machine as an abstract program ("der Maschine, als Symbol" §193) and the actual physical machine, which is subject to breakdown ("do we forget the possibility of their bending, breaking off, melting, and so on?" (§193)). The dispositional theory views the subject himself as a kind of machine, whose potential actions embody the function. So in this sense the dispositional theory and the idea of the machine-as-embodying-the-function are really one. Wittgenstein's attitude toward both is the same: they confuse the 'hardness of a rule' with the 'hardness of a material' (*RFM*, II [III], §87). On my interpretation, then, Wittgenstein agrees with his interlocutor (§194 and §195) that the sense in which all the values of the function are already present is not simply causal, although he disagrees with the idea that the future use is already present in some mysterious non-causal way.

The finite automaton: no malfunction and an infinite life.

Although, in an attempt to follow Wittgenstein, I have emphasized the distinction between concrete physical machines and their abstract programs in what I have written above, it might be instructive to look at the outcome when the limitation of machines is idealized as in the modern theory of automata. A finite automaton, as usually defined, has only finitely many states, receives only finitely many distinct inputs, and has only finitely many outputs, but it is idealized in two respects: it has no problem of malfunction, and its lifetime (without any decay or wearing out of its parts) is infinite. Such a machine can, in a sense, perform computations on arbitrarily large whole numbers. If it has notations for

The automaton adds digit by digit but provably can't multiply.

the single digits from zero through nine, inclusive, it can receive arbitrarily large positive whole numbers as inputs simply by being given their digits one by one. (We cannot do this, since our effective lifetimes are finite, and there is a minimum time needed for us to understand any single digit.) Such an automaton can add according to the usual algorithm in decimal notation (the digits for the numbers being added should be fed into the machine starting from the last digits of both summands and going backwards, as in the usual algorithm). However, it can be proved that, in the same ordinary decimal notation, such a machine cannot *multiply*. Any function computed by such a machine that purports to be multiplication will, for large enough arguments, exhibit 'quus-like' (or rather, 'quimes-like') properties at sufficiently large arguments. Even if we were idealized as finite automata, a dispositional theory would yield unacceptable results.

Suppose we idealized even further and considered a Turing machine which has a tape to use which is infinite in both directions. Such a machine has infinite extent at every moment, in addition to an infinite lifetime without malfunctions. Turing machines can multiply correctly, but it is well known that even here there are many functions we can define explicitly that can be computed by no such machine. A crude dispositional theory would attribute to us a non-standard interpretation (or no interpretation at all) for any such function. (See above, note 20.)

I have found that both the crude dispositional theory and the function-as-embodied-in-a-machine come up frequently when Wittgenstein's paradox is discussed. For this reason, and because of their close relation to Wittgenstein's text, I have expounded these theories, though sometimes I have wondered whether the discussion of them is excessively long. On the other hand, I have resisted the temptation to discuss 'functionalism' explicitly, even though various forms of it have been so attractive to so many of the best recent writers that it has almost become the received philosophy of mind in the USA. Especially I have feared that some readers of the discussion in the text will think that 'functionalism' is precisely the way to modify the crude dispositional theory so as to meet the criticisms (especially those that rely on the circularity of *ceteris paribus* clauses). (I report, however, that thus far I have not run into such reactions in practice.) I cannot discuss functionalism at length here without straying from the main point. But I offer a brief hint. Functionalists are fond of comparing psychological states to the abstract states of a (Turing) machine, though some are cognizant of certain limitations of the comparison. All regard psychology as given by a set of causal connections, analogous to the *causal* operation of a machine. But then the remarks of the text stand here as well: any concrete physical object can be viewed as an imperfect realization of many machine programs. Taking a human organism as a concrete object, what is to tell

The moral: meaning's relation to action is normative, not descriptive.

The moral of the present discussion of the dispositional account may be relevant to other areas of concern to philosophers beyond the immediate point at issue. Suppose I do mean addition by '+'. What is the relation of this supposition to the question how I will respond to the problem '68+57'? The dispositionalist gives a *descriptive* account of this relation: if '+' meant addition, then I will answer '125'. But this is not the proper account of the relation, which is *normative*, not descriptive. The point is *not* that, if I meant addition by '+', I *will* answer '125', but that, if I intend to accord with my past meaning of '+', I *should* answer '125'. Computational error, finiteness of my capacity, and other disturbing factors may lead me not to be *disposed* to respond as I *should*, but if so, I have not acted in accordance with my intentions. The relation of meaning and intention to future action is *normative*, not *descriptive*.

Full circle: the justificatory role is what dispositions ignore.

In the beginning of our discussion of the dispositional analysis, we suggested that it had a certain air of irrelevance with respect to a significant aspect of the sceptical problem – that the fact that the sceptic can maintain the hypothesis that I meant quus shows that I had no *justification* for answering '125' rather than '5'. How does the dispositional analysis even appear to touch this problem? Our conclusion in the previous paragraph shows that in some sense, after giving a number of more specific criticisms of the dispositional theory, we have returned full circle to our original intuition. Precisely the fact that our answer to the question of which function I meant is *justificatory* of my present response is ignored in the dispositional account and leads to all its difficulties.

I shall leave the dispositional view. Perhaps I have already belabored it too much. Let us repudiate briefly another

---

us *which* program he should be regarded as instantiating? In particular, does he compute 'plus' or 'quus'? If the remarks on machines in my own (and Wittgenstein's) text are understood, I think it will emerge that as far as the present problem is concerned, Wittgenstein would regard his remarks on machines as applicable to 'functionalism' as well.

I hope to elaborate on these remarks elsewhere.

Simplicity: the appeal rests on misreading the problem.

suggestion. Let no one – under the influence of too much philosophy of science – suggest that the hypothesis that I meant plus is to be preferred as the *simplest* hypothesis. I will not here argue that simplicity is relative, or that it is hard to define, or that a Martian might find the quus function simpler than the plus function. Such replies may have considerable merit, but the real trouble with the appeal to simplicity is more basic. Such an appeal must be based either on a misunderstanding of the sceptical problem, or of the role of simplicity considerations, or both. Recall that the sceptical problem was not merely epistemic. The sceptic argues that there is no fact as to what I meant, whether plus or quus. Now simplicity considerations can help us decide between competing hypotheses, but they obviously can never tell us what the competing hypotheses are. If we do not understand what two hypotheses *state*, what does it mean to say that one is 'more probable' because it is 'simpler'? If the two competing hypotheses are not genuine hypotheses, not assertions of genuine matters of fact, no 'simplicity' considerations will make them so.

Simplicity picks between hypotheses, it can't make them genuine.

The electron analogy: a being who sees the facts needs no simplicity.

Suppose there are two conflicting hypotheses about electrons, both confirmed by the experimental data. If our own view of statements about electrons is 'realist' and not 'instrumentalist', we will view these assertions as making factual assertions about some 'reality' about electrons. God, or some appropriate being who could 'see' the facts about electrons directly, would have no need for experimental evidence or simplicity considerations to decide between hypotheses. We, who lack such capacities, must rely on indirect evidence, from the effects of the electrons on the behavior of gross objects, to decide between the hypotheses. If two competing hypotheses are indistinguishable as far as their effects on gross objects are concerned, then *we* must fall back on simplicity considerations to decide between them. A being – not ourselves – who could 'see' the facts about electrons 'directly' would have no need to invoke simplicity considerations, nor to rely on indirect evidence to decide between the hypotheses; he would 'directly perceive' the relevant facts that

make one hypothesis true rather than another. To say this is simply to repeat, in colorful terminology, the assertion that the two hypotheses do state genuinely different matters of fact.

Not epistemic scepticism, so an omniscient being needs no simplicity.

Now Wittgenstein's sceptic argues that he knows of no fact about an individual that could constitute his state of meaning plus rather than quus. Against *this* claim simplicity considerations are irrelevant. Simplicity considerations would have been relevant against a sceptic who argued that the indirectness of our access to the facts of meaning and intention *prevents us ever from knowing* whether we mean plus or quus. But such merely epistemological scepticism is *not* in question. The sceptic does not argue that our own limitations of access to the facts prevent us from knowing something hidden. He claims that an omniscient being, with access to *all* available facts, still would not find any fact that differentiates between the plus and the quus hypotheses. Such an omniscient being would have neither need nor use for simplicity considerations.[25]

[25] A different use of 'simplicity', not that by which we evaluate competing theories, might suggest itself with respect to the discussion of machines above. There I remarked that a concrete physical machine, considered as an object without reference to a designer, may (approximately) instantiate any number of programs that (approximately, allowing for some 'malfunctioning') extend its actual finite behavior. If the physical machine was not designed but, so to speak, 'fell from the sky', there can be no fact of the matter as to which program it 'really' instantiates, hence no 'simplest hypothesis' about this non-existent fact.

Nevertheless, given a physical machine, one might ask what is the *simplest program* that the physical machine approximates. To do this one would have to find a measure of the simplicity of programs, a measure of the trade-off of the simplicity of the program with the degree to which the concrete machine fails to conform to it (malfunctions), and so on. I who am no expert, nor even an amateur, am unaware that this problem has been considered by theoretical computer scientists. Whether or not it has been considered, intuition suggests that something might be made of it, though it would not be trivial to find simplicity measures that give intuitively satisfying results.

I doubt that any of this would illuminate Wittgenstein's sceptical paradox. One might try, say, to define the function I meant as the one that, according to the simplicity measure, followed the simplest program

Don't I know directly that I mean plus? No hypothetical state will do.

The idea that we lack 'direct' access to the facts whether we mean plus or quus is bizarre in any case. Do I not know, directly, and with a fair degree of certainty, that I mean plus? Recall that a fact as to what I mean now is supposed to *justify* my future actions, to make them *inevitable* if I wish to use words with the same meaning with which I used them before. This was our fundamental requirement on a fact as to what I meant. No 'hypothetical' state could satisfy such a requirement: If I can only form hypotheses as to whether I now mean plus or quus, if the truth of the matter is buried deep in my unconscious and can only be posited as a tentative hypothesis, then in the future I can only proceed hestitatingly and hypothetically, *conjecturing* that I probably ought to answer '68+57' with '125' rather than '5'. Obviously, this is not an accurate account of the matter. There may be some facts about me to which my access is indirect, and about which I must form tentative hypotheses: but surely the fact as to what I mean by 'plus' is not one of them! To say that it is, is already to take a big step in the direction of scepticism. Remember that I immediately and unhesitatingly calculate '68 + 57'as I do, and the meaning I assign to '+' is supposed to *justify* this procedure. I do not form tentative hypotheses, wondering what I should do if one hypothesis or another were true.

Treating the fact as hidden is already a step toward scepticism.

Now the reference, in our exposition, to what an omniscient being could or would know is merely a dramatic device. When the sceptic denies that even God, who knows all the

---

approximately compatible with my physical structure. Suppose brain physiologists found – to their surprise – that actually such a simplicity measure led to a program that did not compute addition for the '+' function, but some other function. Would this show that I did not mean addition by '+'? Yet, in the absence of detailed knowledge of the brain (and the hypothetical simplicity measure), the physiological discovery in question is by no means inconceivable. The justificatory aspect of the sceptic's problem is even more obviously remote from any such simplicity measure. I do not justify my choice of '125' rather than '5' as an answer to '68+57' by citing a hypothetical simplicity measure of the type mentioned. (I hope to elaborate on this in the projected work on functionalism mentioned in note 24 above.)

God as a dramatic device, and the demand for a reduction.

facts, could know whether I meant plus or quus, he is simply giving colorful expression to his denial that there is any fact of the matter as to which I meant. Perhaps if we remove the metaphor we may do better. The metaphor, perhaps, may seduce us towards scepticism by encouraging us to look for a reduction of the notions of meaning and intention to something else. Why not argue that "meaning addition by 'plus'" denotes an irreducible experience, with its own special *quale*, known directly to each of us by introspection? (Headaches, tickles, nausea are examples of inner states with such *qualia*.)[26] Perhaps the "decisive move in the conjuring trick" has been made when the sceptic notes that I have performed only finitely many additions and challenges me, in the light of *this* fact, to adduce some fact that 'shows' that I did not mean quus. Maybe I appear to be unable to reply just because the experience of meaning addition by 'plus' is as unique and irreducible as that of seeing yellow or feeling a headache, while the sceptic's challenge invites me to look for another fact or experience to which this can be reduced.

The conjuring trick: was it the demand to reduce meaning?

I referred to an *introspectible* experience because, since each of us knows immediately and with fair certainty that he means addition by 'plus', presumably the view in question assumes we know this in the same way we know that we have headaches – by attending to the 'qualitative' character of our own experiences. Presumably the experience of *meaning addition* has its own irreducible quality, as does that of feeling a headache. The fact that I mean addition by 'plus' is to be identified with my possession of an experience of this quality.

Meaning addition as a quale known like a headache.

Once again, as in the case of the dispositional account, the proffered theory seems to be off target as an answer to the original challenge of the sceptic. The sceptic wanted to know why I was so sure that I ought to say '125', when asked about '68+57'. I had never thought of this particular addition before: is not an interpretation of the '+' sign as quus compatible with everything I thought? Well, suppose I do in fact feel a certain

Off target again: suppose I do feel a special headache.

[26] It is well known that this type of view is characteristic of Hume's philosophy. See note 51 below.

The headache can't tell me whether to say 125 or 5.

headache with a very special quality whenever I think of the '+' sign. How on earth would this headache help me figure out whether I ought to answer '125' or '5' when asked about '68+57'? If I think the headache indicates that I ought to say '125', would there be anything about it to refute a sceptic's contention that, on the contrary, it indicates that I should say '5'? The idea that each of my inner states – including, presumably, meaning what I do by 'plus' – has its special discernible quality like a headache, a tickle, or the experience of a blue after-image, is indeed one of the cornerstones of classical empiricism. Cornerstone it may be, but it is very hard to see how the alleged introspectible *quale* could be relevant to the problem at hand.

An image fares no better: the drawing of a cube, §139.

Similar remarks apply even to those cases where the classical empiricist picture might seem to have a greater plausibility. This picture suggested that association of an image with a word (paradigmatically a visual one) determined its meaning. For example (§139), a drawing of a cube comes to my mind whenever I hear or say the word 'cube'. It should be obvious that this need not be the case. Many of us use words such as 'cube' even though no such drawing or image comes to mind. Let us suppose, however, for the moment that one does. 'In what sense can this picture fit or fail to fit a use of the word "cube"? – Perhaps you say: "It's quite simple; – if that picture comes to me and I point to a triangular prism for instance, and say it's a cube, then this use of the word doesn't fit the picture." But doesn't it fit? I have purposely so chosen the example that it is quite easy to imagine a *method of projection* according to which the picture does fit after all. The picture of the cube did indeed *suggest* a certain use to us, but it was possible for me to use it differently.' The sceptic could suggest that the image be used in non-standard ways. 'Suppose, however, that not merely the picture of the cube, but also the method of projection comes before our mind? – How am I to imagine this? – Perhaps I see before me a schema showing the method of projection: say a picture of two cubes connected by lines of projection. – But does this really get me any further?

The schema of projection can be reinterpreted too.

No impression with a quale, nor a pile of them, says how to apply it.

Can't I now imagine different applications of this schema too?' (§141). Once again, a rule for interpreting a rule. No internal impression, with a *quale*, could possibly tell me in itself how it is to be applied in future cases. Nor can any pile up of such impressions, thought of as rules for interpreting rules, do the job.[27] The answer to the sceptic's problem, "What tells me how I am to apply a given rule in a new case?", must come from something outside any images or 'qualitative' mental states. This is obvious, in the case of 'plus' – it is clear enough that no internal state such as a headache, a tickle, an image, could do the job. (Obviously I do not have an image of the infinite table of the 'plus' function in my mind. Some such image would be the only candidate that even has surface plausibility as a device for telling me how to apply 'plus'.) It may be less obvious in other cases, such as 'cube', but in fact it is also true of such cases as well.

Two claims: the experience wouldn't help, and it doesn't exist.

So: If there were a special experience of 'meaning' addition by 'plus', analogous to a headache, it would not have the properties that a state of meaning addition by 'plus' ought to have – it would not tell me what to do in new cases. In fact, however, Wittgenstein extensively argues in addition that the supposed unique special experience of meaning (addition by 'plus', etc.) does not exist. His investigation here is an introspective one, designed to show that the supposed unique experience is a chimera. Of all the replies to the sceptic he combats, the view of meaning as an introspectible experience is probably the most natural and fundamental. But for the present day audience I dealt with it neither first nor at greatest length, for, though the Humean picture of an irreducible 'impression' corresponding to each psychological state or event has tempted many in the past, it tempts relatively few today. In fact, if in the past it was too readily and simplistically assumed, at present its force is – at least in my personal opinion – probably too *little* felt. There are several reasons for this. One is that, in this instance, Wittgenstein's critique of alternative

The most natural reply, though it tempts few today.

[27] The remarks above, p. 20, on the use of an image, or even a physical sample, of green make the same point.

views has been relatively well received and absorbed. And related writers – such as Ryle – have reinforced the critique of the Cartesian and Humean pictures. Another reason – unattractive to the present writer – has been the popularity of materialistic-behavioristic views that ignore the problem of felt qualities of mental states altogether, or at least attempt to analyze all such states away in broadly behavioristic terms.[28]

Not a behaviourist premise: the case is made by introspection.

It is important to repeat in the present connection what I have said above: Wittgenstein does not base his considerations on any behavioristic *premise* that dismisses the 'inner'. On the contrary, much of his argumentation consists in detailed introspective considerations. Careful consideration of our inner lives, he argues, will show that there is no special inner experience of 'meaning' of the kind supposed by his opponent. The case is specifically in *contrast* with feeling a pain, seeing red, and the like.

Attend to when I learned to add: was there a Eureka moment?

It takes relatively little introspective acuteness to realize the dubiousness of the attribution of a special qualitative character to the 'experience' of meaning addition by 'plus'. Attend to what happened when I first learned to add. First, there may or may not have been a specifiable time, probably in my childhood, at which I suddenly felt (*Eureka!*) that I had grasped the rule for addition. If there was not, it is very hard to see in what the suppositious special experience of my learning to add consisted. Even if there was a particular time at which I could have shouted "*Eureka!*" – surely the exceptional case – in what did the attendant experience consist? Probably consideration of a few particular cases and a thought – "Now I've got it!" – or the like. Could just *this* be the content of an experience of 'meaning addition'? How would it have been different if I had

[28] Although there are clear classical senses of behaviorism in which such current philosophies of mind as 'functionalism' are not behaviorist, nevertheless, speaking for myself, I find much contemporary 'functionalism' (especially those versions that attempt to give 'functional' *analyses* of mental terms) are far too behavioristic for my own taste. It would require an extensive digression to go into the matter further here.

Add 5+7 now: would quadding or multiplying have felt different?

meant quus? Suppose I perform a particular addition now, say '5+7'. Is there any special quality to the experience? Would it have been different if I had been trained in, and performed, the corresponding quaddition? How different indeed would the *experience* have been if I had performed the corresponding multiplication ('5×7'), other than that I would have responded automatically with a different answer? (Try the experiment yourself.)

Wittgenstein returns to points like these repeatedly throughout *Philosophical Investigations*. In the sections where he discusses his sceptical paradox (§§137–242), after a general consideration of the alleged introspectible process of understanding, he considers the issue in connection with the special case of *reading* (§§156–78). By 'reading' Wittgenstein means reading out loud what is written or printed and similar activities: he is not concerned with understanding what is written. I myself, like many of my coreligionists, first learned to 'read' Hebrew in this sense before I could understand more than a few words of the language. Reading in this sense is a simple case of 'following a rule'. Wittgenstein points out that a beginner, who reads by laboriously spelling words out, may have an introspectible experience when he really reads, as opposed to pretending to 'read' a passage he has actually memorized in advance; but an experienced reader simply calls the words out and is aware of no special conscious experience of 'deriving' the words from the page. The experienced reader may 'feel' nothing different when he reads from what the beginner feels, or does not feel, when he pretends. And suppose a teacher is teaching a number of beginners to read. Some pretend, others occasionally get it right by accident, others have already learned to read. When has someone passed into the latter class? In general, there will not be an identifiable moment when this has happened: the teacher will judge of a given pupil that he has 'learned to read' if he passes tests for reading often enough. There may or may not be an identifiable moment when the pupil first *felt*, "Now I am reading!" but the

presence of such an experience is neither a necessary nor a sufficient condition for the teacher to judge of him that he is reading.

Again (§160), someone may, under the influence of a drug, or in a dream, be presented with a made-up 'alphabet' and utter certain words, with all the characteristic 'feeling' of reading, to the extent that such a 'feeling' exists at all. If, after the drug wears off (or he wakes up), he himself thinks he was uttering words at random with no real connection with the script, should we really say he was reading? Or, on the other hand, what if the drug leads him to read fluently from a genuine text, but with the 'sensation' of reciting something learned by heart? Wasn't he still reading?

It is by examples like these – *Philosophical Investigations* contains a wealth of examples and mental thought experiments beyond what I have summarized – that Wittgenstein argues that the supposed special 'experiences' associated with rule following are chimerical.[29] As I said, my own discussion

[29] The point should not be overstated. Although Wittgenstein does deny that there is any particular 'qualitative' experience like a headache, present when and only when we use a word with a certain meaning (or read, or understand, etc.), he does acknowledge a certain 'feel' to our meaningful use of a word that may under certain circumstances be lost. Many have had a fairly common experience: by repeating a word or phrase again and again, one may be able to deprive it of its normal 'life', so that it comes to sound strange and foreign, even though one is still able to utter it under the right circumstances. Here there is a special feeling of foreignness in a particular case. Could there be someone who always used words like a mechanism, without any 'feeling' of a distinction between this mechanistic type of use and the normal case? Wittgenstein is concerned with these matters in the second part of the *Investigations*, in connection with his discussion of 'seeing as' (section xi, pp. 193–229). Consider especially his remarks on 'aspect blindness', pp. 213–14, and the relation of 'seeing an aspect' to 'experiencing the meaning of a word', p. 214. (See his examples on p. 214: "What would you be missing . . . if you did not feel that a word lost its meaning and became a mere sound if it was repeated ten times over? . . . Suppose I had agreed on a code with someone; "tower" means bank. I tell him "Now go to the tower" – he understands me and acts accordingly, but he feels the word "tower" to be strange in this use, it has not yet 'taken on' the meaning." He gives many examples on pp. 213–18.)

Compare (as Wittgenstein does) the feeling of meaning a word as such-and-such (think of 'till' now as a verb, now as a noun, etc.), with the idea of visual aspects discussed at length in section xi of the second part of the *Investigations*. We can see the duck-rabbit (p. 194) now as a rabbit, now as a duck; we can see the Necker cube, now with one face forward, now with another; we can see a cube drawing (p. 193) as a box, a wire frame, etc. How, if at all, does our visual experience change? The experience is much more elusive than is anything like the feeling of a headache, the hearing of a sound, the visual experience of a blue patch. The corresponding 'aspects' of meaning would seem to be introspectively even more elusive.

Similarly, although some of the passages in §§156–78 seem to debunk the idea of a conscious special experience of 'being guided' (when reading) altogether, it seems wrong to think of it as totally dismissed. For example, in §160, Wittgenstein speaks both of the 'sensation of saying something he has learnt by heart' and of the 'sensation of reading', though the point of the paragraph is that the presence or absence of such sensations is not what constitutes the distinction between reading, saying something by heart, and yet something else. To some extent, I think Wittgenstein's discussion may have a certain ambivalence. Nevertheless, some relevant points made are these: (i) Whatever an 'experience of being guided' (in reading) may be, it is not something with a gross and introspectible qualitative character, like a headache (contrary to Hume). (ii) In particular cases of reading, we may feel definite and introspectible experiences, but these are different and distinct experiences, peculiar to each individual case, not a single experience present in all cases. (In the same way, Wittgenstein speaks of various introspectible 'mental processes' that *in particular circumstances* occur when I understand a word – see §§151–5, but none of these *is* the 'process' of understanding, indeed understanding is not a 'mental process' – see pp. 49–51 below. The discussion of reading, which follows §§151–5 immediately, is meant to illustrate these points. (iii) Perhaps most important, whatever the elusive feeling of being guided may be, its presence or absence is not constitutive of whether I am reading or not. See, for example, the cases mentioned above in the text, of the pupil learning to read and of the person under the influence of a drug.

Rush Rhees, in his preface to *The Blue and Brown Books* (Basil Blackwell, Oxford and Harper and Brothers, New York, 1958, xiv+185 pp.) emphasizes (see pp. xii–xiv) the problem created for Wittgenstein by 'meaning blindness', and he emphasizes that the discussion of 'seeing something as something' in section xi of the second

can be brief because this particular Wittgensteinian lesson has been relatively well learned, perhaps too well learned. But some points should be noted. First, to repeat, the method of the investigation, and of the thought-experiments is deeply introspective: it is exactly the kind of investigation a strict psychological behaviorist would *prohibit*.[30] Second, although Wittgenstein does conclude that behavior, and dispositions to behavior, lead us to *say* of a person that he is reading, or adding, or whatever, this should not, in my opinion, be misconstrued as an endorsement of the dispositional theory: he does not say that reading or adding *is* a certain disposition to behavior.[31]

part of *Philosophical Investigations* is motivated by an attempt to deal with the elusive question. Earlier portions of the *Investigations* repudiate traditional pictures of internal, qualitative states of meaning and understanding; but later Wittgenstein seems, as Rhees says, to be worried that he may be in danger of replacing the classical picture by an overly mechanistic one, though certainly he still repudiates any idea that a certain qualitative experience *is* what constitutes my using words with a certain meaning. Could there be a 'meaning blind' person who operated with words just as we do? If so, would we say that he is as much in command of the language as we? The 'official' answer to the second question, as given in our main text, is 'yes'; but perhaps the answer should be, "Say what you want, as long as you know the facts." It is not clear that the problem is entirely resolved. Note that here, too, the discussion is introspective, based on an investigation of our own phenomenal experience. It is not the kind of investigation that would be undertaken by a behaviorist. No doubt the matter deserves a careful and extended treatment.

[30] §314 says: "It shows a fundamental misunderstanding, if I am inclined to study the headache I have now in order to get clear about the fundamental philosophical problem of sensation." If this remark is to be consistent with Wittgenstein's frequent practice as outlined in the text above and note 29, it *cannot* be read as *generally* condemning the philosophical use of introspective reflections on the phenomenology of our experience.

[31] I should not deny that Wittgenstein has important affinities to behaviorism (as to finitism — see pp. 105–7 below). Such a famous slogan as "My attitude towards him is an attitude towards a soul (*Seele*). I am not of the *opinion* that he has a soul" (p. 178) sounds much too behavioristic for me. I personally would like to think that anyone who does not think of me as conscious is wrong about the facts, not simply 'unfortunate', or 'evil', or

Wittgenstein's conviction of the contrast between states of understanding, reading and the like, and 'genuine', introspectible mental states or processes is so strong that it leads him – who is often regarded as a (or the) father of 'ordinary language philosophy', and who emphasizes the importance of respect for the way language is actually used – into some curious remarks about ordinary usage. Consider §154: "In the sense in which there are processes (including mental processes) which are characteristic of understanding, understanding is not a mental process. (A pain's growing more and less; the hearing of a tune or sentence: these are mental processes.)" Or again, at the bottom of p. 59, "'Understanding a word': a state. But a *mental* state? – Depression, excitement, pain, are called mental states. Carry out a grammatical investigation . . ." The terms 'mental state' and 'mental process' have a somewhat theoretical flavor, and I am not sure how firmly one can speak of their 'ordinary' use. However, my own linguistic intuitions do not entirely agree with Wittgenstein's remarks.[32] Coming to understand, or learning, seems to me to

even 'monstrous' or 'inhuman', in his 'attitude' (whatever that might mean).

(If '*Seele*' is translated as 'soul', it might be thought that the 'attitude' (*Einstellung*') to which Wittgenstein refers has special religious connotations, or associations with Greek metaphysics and the accompanying philosophical tradition. But it is clear from the entire passage that the issue relates simply to the difference between my 'attitude' toward a conscious being and toward an automaton, even though one of the paragraphs refers specifically to the religious doctrine of the immortality of the soul ('*Seele*'). Perhaps in some respects 'mind' might be a less misleading translation of '*Seele*' in the sentence quoted above, since for the contemporary English speaking philosophical reader it is somewhat less loaded with special philosophical and religious connotations. I feel that this may be so even if 'soul' captures the flavor of the German '*Seele*' better than 'mind'. Anscombe translates '*Seele*' and its derivatives sometimes as 'soul', sometimes as 'mind', depending on the context. The problem really seems to be that German has only '*Seele*' and '*Geist*' to do duty where an English speaking philosopher would use 'mind'. See also the postscript below, note 11.

[32] These are my intuitions in English. I have no idea whether any differences

be a 'mental process' if anything is. A pain's growing more and less, and especially the hearing of a tune or sentence, are probably not ordinarily thought of as 'mental' processes at all. Although depression and anxiety would ordinarily be called 'mental' states, pain (if genuine physical pain is meant) is probably *not* a 'mental' state. ("It's all in your mind" means that no genuine physical pain is present.) But Wittgenstein's concern is not really with usage but with a philosophical terminology. 'Mental states' and 'mental processes' are those introspectible 'inner' contents that I can find in my mind, or that God could find if he looked into my mind.[33] Such

with the German ('*seelischer Vorgang*' and '*seelischer Zustand*'), in nuance or usage, affect the matter.

[33] Or so it would seem from the passages quoted. But the denial that understanding is a 'mental process' in §154 is preceded by the weaker remark, "Try not to think of understanding as a 'mental process' at all – for that is the expression which confuses you." In itself, this seems to say that thinking of understanding as a 'mental process' leads to misleading philosophical pictures, but not necessarily that it is wrong. See also §§305–6: ' "But you surely cannot deny that, for example, in remembering, an inner process takes place."—What gives the impression that we want to deny anything? . . . What we deny is that the picture of the inner process gives us the correct use of the word "to remember" . . . Why should I deny that there is a mental process? But "There has just taken place in me the mental process of remembering . . ." means nothing more than: "I have just remembered . . ." To deny the mental process would mean to deny the remembering; to deny that anyone ever remembers anything.' *This* passage gives the impression that *of course* remembering is a 'mental process' if anything is, but that this ordinary terminology is philosophically misleading. (The German here is *'geistiger Vorgang'* while in the earlier passages it was *'seelischer Vorgang'* (§154) and *'seelischer Zustand'* (p. 59), but as far as I can see, this has no significance beyond stylistic variation. It is possible that the fact that Wittgenstein speaks here of remembering, while earlier he had spoken of understanding, is significant, but even this seems to me to be unlikely. Note that in §154, the genuine 'mental processes' are a pain's growing more or less, the hearing of a tune or sentence – processes with an 'introspectible quality' in the sense we have used the phrase. For Wittgenstein remembering is not a process like these, even though, as in the case of understanding in §154, there may be processes with introspectible qualities that take place when we remember. Assuming that the examples given in §154 are meant to be typical 'mental processes', the examples would be very misleading unless remembering were taken not to be a 'mental process' in the sense of §154. Remembering, like understanding, is an 'intentional' state (see note 19 above) subject to Wittgenstein's sceptical problem.) See also the discussion of 'incorporeal processes' in §339.)

phenomena, inasmuch as they are introspectible, 'qualitative' states of the mind, are not subject to immediate sceptical challenge of the present type. Understanding is not one of these.

The desperate option: a primitive state of a unique kind.

Of course the falsity of the 'unique introspectible state' view of meaning plus must have been implicit from the start of the problem. If there really were an introspectible state, like a headache, of meaning addition by 'plus' (and if it really could have the justificatory role such a state ought to have), it would have stared one in the face and would have robbed the sceptic's challenge of any appeal. But given the force of this challenge, the need philosophers have felt to posit such a state and the loss we incur when we are robbed of it should be apparent. Perhaps we may try to recoup, by arguing that meaning addition by 'plus' is a state even more *sui generis* than we have argued before. Perhaps it is simply a primitive state, not to be assimilated to sensations or headaches or any 'qualitative' states, nor to be assimilated to dispositions, but a state of a unique kind of its own.

Irrefutable maybe, but mysterious, and then the logical difficulty.

Such a move may in a sense be irrefutable, and if it is taken in an appropriate way Wittgenstein may even accept it. But it seems desperate: it leaves the nature of this postulated primitive state – the primitive state of 'meaning addition by "plus"' – completely mysterious. It is not supposed to be an introspectible state, yet we supposedly are aware of it with some fair degree of certainty whenever it occurs. For how else can each of us be confident that he *does*, at present, mean addition by 'plus'? Even more important is the logical difficulty implicit in Wittgenstein's sceptical argument. I think that Wittgenstein argues, not merely as we have said hitherto, that introspection shows that the alleged 'qualitative' state of understanding is a

chimera, but also that it is logically impossible (or at least that there is a considerable logical difficulty) for there to be a state of 'meaning addition by "plus"' at all.

A finite state with every case somehow present in it, §195.

Such a state would have to be a finite object, contained in our finite minds.[34] It does not consist in my explicitly thinking of each case of the addition table, nor even of my encoding each separate case in the brain: we lack the capacity for that. Yet (§195) "in a *queer* way" each such case already is "in some sense present". (Before we hear Wittgenstein's sceptical argument, we surely suppose – unreflectively – that something like this is indeed the case. Even now I have a strong inclination to think this somehow must be right.) What can that sense be? Can we conceive of a finite state which *could* not be interpreted in a quus-like way? How could that be? The proposal I am now discussing brushes such questions under the rug, since the nature of the supposed 'state' is left

[34] We have stressed that I think of only finitely many cases of the addition table. Anyone who claims to have thought of infinitely many cases of the table is a liar. (Some philosophers – probably Wittgenstein – go so far as to say that they see a conceptual incoherence in the supposition that anyone thought of infinitely many such cases. We need not discuss the merits of this strong view here as long as we acknowledge the weaker claim that as a matter of fact each of us thinks of only finitely many cases.) It is worth noting, however, that although it is useful, following Wittgenstein himself, to *begin* the presentation of the puzzle with the observation that I have thought of only finitely many cases, it appears that in principle this particular ladder can be kicked away. Suppose that I had explicitly thought of *all* cases of the addition table. How can this help me answer the question '68+57'? Well, looking back over my own mental records, I find that I gave myself explicit directions. "If you are ever asked about '68+57', reply '125'!" Can't the sceptic say that these directions, too, are to be interpreted in a non-standard way? (See *Remarks on the Foundations of Mathematics*, I, §3: "If I know it *in advance*, what use is this knowledge to me later on? I mean: how do I know what to do with this earlier knowledge when the step is actually taken?") It would appear that, if finiteness is relevant, it comes more crucially in the fact that "justifications must come to an end somewhere" than in the fact that I think of only finitely many cases of the addition table, even though Wittgenstein stresses both facts. Either fact can be used to develop the sceptical paradox; both are important.

Footnote 34: even an infinite table could be reinterpreted.

Finitude matters as justifications ending, not as finite cases.

§195 quoted in full: not causal, yet the use is somehow present.

mysterious. "But" – to quote the protest in §195 more fully – "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." A causal determination is the kind of analysis supposed by the dispositional theorist, and we have already seen that that is to be rejected. Presumably the relation now in question grounds some entailment roughly like: "If I now mean addition by 'plus'; then, if I remember this meaning in the future and wish to accord with what I meant, and do not miscalculate, then when asked for '68+57', I will respond '125'." If Hume is right, of course, no past state of my mind can entail that I will give any particular response in the future. But that I meant 125 in the past does not itself entail this; I must remember what I meant, and so on. Nevertheless it remains mysterious exactly how the existence of *any* finite past state of my mind could entail that, if I wish to accord with it, and remember the state, and do not miscalculate, I must give a determinate answer to an arbitrarily large addition problem.[35]

How could a finite past state entail an answer to any sum?

Platonism: the function contains its triples, a fact of mathematics.

Mathematical realists, or 'Platonists', have emphasized the non-mental nature of mathematical entities. The addition function is not in any particular mind, nor is it the common property of all minds. It has an independent, 'objective', existence. There is then no problem – as far as the present considerations go – as to how the addition function (taken, say, as a set of triples)[36] contains within it all its instances, such as the triple (68, 57, 125). This simply is in the nature of the mathematical object in question, and it may well be an infinite

[35] See p. 218: "Meaning it is not a process which accompanies a word. For no process could have the consequences of meaning." This aphorism makes the general point sketched in the text. No process can entail what meaning entails. In particular, no process could entail the rough conditional stated above. See the discussion below, pp. 93–4, of Wittgenstein's view of these conditionals.

[36] Of course Frege would not accept the identification of a function with a set of triples. Such an identification violates his conception of functions as 'unsaturated'. Although this complication is very important for Frege's philosophy, it can be ignored for the purposes of the present presentation.

object. The proof that the addition function contains such a triple as (68, 57, 125) belongs to mathematics and has nothing to do with meaning or intention.

Frege's four elements: function, sign, sense and idea.

Frege's analysis of the usage of the plus sign by an individual posits the following four elements: (a) the addition function, an 'objective' mathematical entity; (b) the addition sign '+', a linguistic entity; (c) the 'sense' of this sign, an 'objective' abstract entity like the function; (d) an idea in the individual's mind associated with the sign. The idea is a 'subjective' mental entity, private to each individual and different in different minds. The 'sense', in contrast, is the same for all individuals who use '+' in the standard way. Each such individual grasps this sense by virtue of having an appropriate idea in his mind. The 'sense' in turn *determines* the addition function as the *referent* of the '+' sign.

A sense determines its referent by nature, but what grasps the sense?

There is again no special problem, for this position, as to the relation between the sense and the referent it determines. It simply is in the nature of a sense to determine a referent. But ultimately the sceptical problem cannot be evaded, and it arises precisely in the question how the existence in my mind of any mental entity or idea can *constitute* 'grasping' any particular sense rather than another. The idea in my mind is a finite object: can it not be interpreted as determining a quus function, rather than a plus function? Of course there may be another idea in my mind, which is supposed to constitute its act of *assigning* a particular interpretation to the first idea; but then the problem obviously arises again at this new level. (A rule for interpreting a rule again.) And so on. For Wittgenstein, Platonism is largely an unhelpful evasion of the problem of how our finite minds can give rules that are supposed to apply to an infinity of cases. Platonic objects may be self-interpreting, or rather, they may need no interpretation; but ultimately there must be some mental entity involved that raises the sceptical problem. (This brief discussion of Platonism is meant for those interested in the issue. If it is so brief that you find it obscure, ignore it.)

Platonism as an evasion: some mental entity must still interpret.