ture. This is partly because of the startling nature of the conclusion that Quine himself drew: that translation is indeterminate (and, since there is no truth about meaning which is *beyond* an interpreter, so is meaning). As well as references in footnotes, the reader might consult:

N. Chomsky, 'Quine's Empirical Assumptions', in Davidson and Hintikka (1969).
W. V. Quine, 'Reply to Chomsky' (ibid.)
— 'On The Reasons for Indeterminacy of Translation', *Journal of Philosophy* (1970).
R. Kirk, 'Underdetermination of Theory and Indeterminacy of Translation', *Analysis* (1973).
C. Boorse, 'The Origins of the Indeterminacy Thesis', *Journal of Philosophy* (1975).
S. Blackburn, 'The Identity of Propositions', in Blackburn (1975).

The requirement which Dummett has insisted upon, that meaning be something which is manifested in the use made of an expression, has given rise to a large but difficult literature. As well as the article by Craig (chapter 2, n. 14 above), the following might be useful:

C. Wright, *Wittgenstein on the Foundations of Mathematics*, esp. Pt. 2.
P. Strawson, 'Scruton and Wright on Anti-Realism', *Proceedings of the Aristotelian Society* (1976).
C. McGinn, 'Truth and Use', in Platts (1980).

There are references to Dummett's revisionism about logic under 6.4 below. The next chapter, which relates meaning to dispositions, is also relevant.

CHAPTER 3

How is Meaning Possible?

Frankly, it is not my words I mistrust, but your minds.
                                   Joseph Conrad, *Lord Jim*.

1. *Describing: Three Ways of Being Odd*

In the last chapter I went through some arguments against supposing that we understand words by connecting them with a directly representative "presence", carrying its own lines of projection onto aspects of the world. We discovered the negative point, as I shall refer to it, that no thing could halt the regress of interpretation. Our conception of what it could be to have a thing, including a diagram, model, or picture, present to the mind, allows for different ways in which the thing could be taken or understood, and this defeats the purpose of the theory, which is itself to explain what it is to take signs one way or another. This negative point means that we must cast around for some other way to come at the phenomenon of meaning.

So far the discussion has been extremely general. Now, however, we specify a little more the kinds of word whose understanding concerns us. This chapter centres upon straightforward, simple applications of predicates to things: the description of things as red, blue, buses, heavy, expensive, and so on. We can apply these terms all right! And we know what we mean when we do so. Our ability to follow principles of application for predicates is our ability to use universals: to classify, think, and judge at all. The difficulties philosophers have found in understanding this kind of ability are not entirely easy to feel. Fortunately however, there are modern ways of approaching these problems, developed more or less simultaneously, and independently, by Wittgenstein, Russell, and Nelson Goodman.[1] Together they put immense pressure upon our understanding of what it could be to assign a meaning to a predicate.

[1] References in notes to this chapter.

Imagine, then, a man going about applying some predicate – say 'round' – to some things, and withholding it from others. He gives the appearance of judging that things have some property or other, which he expresses by means of this term. Perhaps he means that these things are round – he might seem to go through the right procedure to make this his judgement: he assesses distances from the edge of the thing to what we should deem the centre; sees whether such things roll evenly, and perhaps does other tests. He seems like a man judging whether things are round. Now we imagine that he comes upon a new thing – a thing which is quite obviously square. He considers, and he says, 'It's round.' He applies the term! Naturally, our first thought is that he has made a mistake – he has said something which he ought not to have said, something false or incorrect. But two other hypotheses are possible. Perhaps he meant something different by the word – some property which was possessed by the initial set of things to which he applied it, but is also possessed by this square thing. In the terms of chapter 1, his actual language is one in which 'round' means something other than – round. Finally, perhaps he means *nothing* by his term. He is just going through a *parade* of making a genuine judgement as children may sometimes parrot adult judgements without really understanding the terms used in them. So we have three kinds of hypothesis: that he makes a mistake, that he means something different, or that he means nothing at all. What determines which hypothesis is true? What kind of fact is it, which is ultimately the fact in virtue of which one of these hypotheses is correct, and the other two not? For we cannot doubt that on many occasions one *is* correct and the other two are not, although which one varies from occasion to occasion. Mostly, of course, we think that the first is true, because we are used to believing that people mean *something* by what they say, and that mistakes are more common than different usages, particularly of a common word like 'round'.

It will be convenient to have titles for the three hypotheses. In the first, the man means what we do by his term, but makes a mistake. He uses it in accordance with the same rule, or, to put the same thing another way, expresses the same concept or same judgement. That indeed is *why* he is wrong, and says something incorrect, false. Call this the *right-rule* view. In the

second, the man takes the term in a surprising way: he uses it in accordance with some deviant rule, meaning that some of the things which he calls 'round' are what we would also call round, but that others are not. Call this the *bent-rule* view. The last case is that of someone who gives an appearance of applying words to things in accordance with some meaning given to them, but is not really doing so. His sounds are just sounds, capable neither of correctness nor incorrectness, because no rule exists to determine whether a particular "application" of the term is right or wrong. Since an utterance cannot be incorrect, neither can it be correct: no judgement is made. Call this wooden utterance of a term, in which no rule exists to determine significance, the *no-rule* view.

It comes naturally to say: the man himself knows which of these hypotheses is true. So whatever fact it is, which makes one of them true and the others not, it is accessible to him. Perhaps he can introspect it, and make us aware of it in the same way that he can make us aware of other mental facts about himself. *He* knows whether he is making a judgement, and which judgement it is. Whereas we, perhaps, might be less good at judging which of these hypotheses is true, and might get it wrong. On occasion this is how it is. For example, a child might show enormous concentration in writing down an apparently random series of numbers. We might suppose that his production of a numeral is a random event and that the series is not being determined by any rule – just as a lunatic might cover pages and pages with formulae in the belief that he is a mathematician doing great calculations, but signify nothing. The child on this hypothesis has in mind no rule determining a series: his writing one numeral after another is wooden. Nothing counts as correct or incorrect. But the child (like the infant Gauss is said to have done) might surprise us. He might explain which series he is expanding, and show us that he is doing it rightly. We change our mind, and admit that there was method in the madness after all. On a given occasion we might not be sure whether this was going to happen or not. When we don't know, the irresistible image is of something in the child's mind, accessible to him, but only guessed at, perhaps fallibly, by us. We imagine that if we could, as it were, lift off the mental lid – if like God we could look into the glassy essence of the mind – we would then know,

just as the child does, which of the three hypotheses is true.

A great deal of the work of the later Wittgenstein is devoted to showing that this picture of the situation is false. One of his main arguments is the negative point of the last chapter. How do we envisage the subject's own knowledge of his meaning? What does the introspective candidate find as he considers his own mind, which tells him determinately which hypothesis is true? Perhaps pictures, or formulae, or definitions of terms. But the presence of any such thing *cannot be the fact* which determines which hypothesis is true. For any such thing can be taken in different ways. Of course, the presence of an image, or of words framing a definition, or some other presence, might give the candidate confidence that he knows what he means, that one hypothesis out of the first two is right. But it doesn't *make* any particular hypothesis right. It cannot of itself constitute the missing fact, because of the problems of the last chapter, summed up in the negative point. No thing can halt the regress of interpretation, for any thing can be taken in different ways, or in no way at all. Images or words may flit through the candidate's mind, but leave him using the word meaninglessly.

It does not follow that the subject himself is *not* an authority on whether he means anything, or if so, what. All that follows is that we need some different approach to this kind of self-knowledge. It must not be conceived of, as knowledge so often is, as an acquaintance with any kind of presence, mental or otherwise. But that leaves other possibilities. The child in the above example may rightly have perfect confidence that he is genuinely calculating, and that his placing of one number after another in his series is not wooden; he may know this at a point at which an outside observer would not know it. This is not in question: it is the introspective picture of how it can be true which the negative point attacks. Perhaps the candidate knows what he is doing in whichever way we know what we intend or what would please us. The case of the lunatic shows that such confidence can be mistaken, but of course often it is *not* mistaken.

One way in which Wittgenstein pursues the negative point is particularly compelling. He considers someone who understands correctly a simple numerical operation, such as developing a series by adding 2 to the preceding number. We could all

do this, and a learner, after a little instruction with a small sample of such sums, might "catch on" – perhaps in a flash he might come to see what is meant, and then know that he can continue the series correctly, indefinitely. Again, the fact is not in question. And it might seem particularly tempting to think of it in terms of the sudden presence to the mind of a display, a revelation of what is required. But further thought shows that this cannot be right. For consider a later application of the understanding: when the learner writes, say, 188 after 186. If the right-rule view is true, the learner means something which makes it incorrect for him to write anything but this. If he put down 193, he would be wrong. But suppose the second view was true, and the learner had taken our instruction in an unintended, queer way: perhaps he caught on to the bent rule 'add 2 up to 186, and then add 7'. This is a perfectly good instruction – we could programme a computer to follow it, and we might have a purpose in doing so. Now ask: what display in the mental life of someone determined the fact that he took the instructions one way or the other? Not a display of *all* the numbers, because there are too many of them. It would be a pure accident if, in considering the instruction, someone actually thought of *this particular* application. Perhaps a display of some other words: 'Do what was done in this initial sample, whenever *any* number is proposed.' Such a display might occur, of course. But suppose the learner had, in some bent and remarkable fashion, taken the instruction to introduce the bent function 'add 2 up to 186 and then add 7'. Why shouldn't just these words also go through his mind? He could think of himself as "doing what was done in the initial sample whenever any number is proposed". That is, compute *this* function whenever any number is proposed. So the presence of these words does not seem to separate the right-rule learner from a bent-rule learner.

At this point a great variety of issues start to clamour for attention, and it is difficult to preserve a sense of direction. In particular many philosophers see here an opening into a relativistic, conventionalist, view of our own classifications. To *us* it seems absurd and almost incomprehensible that someone would actually take the instructions and the initial sample in this bent way. Why should he get the idea that there is such a gross singularity just at 186? Why didn't he enquire about it?

But might it be that we are here imposing our own, accidental perspective? From the point of view of this bent learner, going on to add 7 after 186 *is* going on in the same way. It is our "similarity space" which, allegedly, he finds bent. Since it came naturally to him to think of the relation between 186 and 193 as like that between 184 and 186, he finds our tendency to insist on 188 as the right successor highly deviant. This conventionalism is sufficiently important to deserve a section to itself. But that is in part a digression from the main issue about meaning. The main issue is not whether there is an element of conventionality in taking the instruction one way or another. The main issue is to obtain some conception of *what it is* to take the instruction in any way at all – in other words, to find out what makes true the right-rule view, rather than the bent-rule view, or the no-rule view. Still, it may help with that problem to think a little further about bent classifications.

2. *Bent Predicates: Wittgenstein and Goodman*

What are we to make of the possibility, if it is one, that somebody takes our instructions, and an initial sample of cases, to introduce a rule of application of some term, but a bent one? It will help to have some examples in mind.

'Add 2'	Bent-rule:	Add 2 up until 186, then add 7.
'Red'	Bent-rule:	A thing is to be called 'red' just if it is observed before 1 January 1986 and is red, or is not so observed, and is yellow.
'round'	Bent-rule:	A thing is to be called 'round' just if it is one of an initial sample, and has the shape defined by a point, travelling equidistantly from a fixed point, or does not belong to the sample, and has four straight sides of equal length, at right-angles to one another.

There is nothing intrinsically wrong with the bent rules. They define perfectly good uses for terms. As I have already remarked, a computer programmed to recognize shapes, if it can determine which things are in the original sample, can give verdicts according to the last rule. Things like canaries which are only observed after 1986 has begun will be properly called 'red' if that term is used in accordance with the given rule – red^{BR} (red-bent-rule), as I shall index it. The bent rules describe meanings which words *can* take: indeed the words used to express them together define them. The curiosity is that terms such as red^{BR} and round^{BR} apply quite properly to the sample of objects which a learner of our ordinary vocabulary will have been shown. So it seems that there could be nothing improper in the learner taking our instructions to introduce these bent meanings. But if he does so, one fine day he will apply the terms in accordance with his understanding to quite astonishing objects – yellow or square objects. In the numerical case he will continue the series in a dramatically devious way.

The bent learner can be thought of graphically in the following way. We can imagine the dimension of colour, shape, or arithmetical addition functions (add 2, add 3, . . .), arranged vertically on the side of a graph. Increase in time (or in number) is plotted along the horizontal axis. We would plot continuities like the line C, and kinks and changes with a line like K in Fig. 2.

Fig. 2

The bent learner has got hold of a "dimension" of properties which reverses the picture. The state of affairs represented to us by C appears to him as a kink; what appears to him as a simple straightforward continuity is the state of affairs which appears bent to us. So his graph is like Fig. 3. Thus, our dimension of arithmetical plus-functions represents someone who adds 2 up to 186 and starts to add 7 after 186 as bent; but the BR

dimension represents him as going on the same way, and somebody who continues to add 2 after that point is represented as kinked. Of course, there is an indefinitely large variety of possible bent dimensions in addition to the examples chosen. They can be created quite automatically.
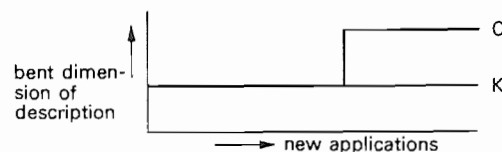

Fig. 3

Goodman used the possibility of these bent dimensions to cast a new light on the classical problem of induction.[2] This problem queries our right to take observed regularities in things as representative – as likely to continue or to have continued in regions of space and time beyond our actual acquaintance. Thus we suppose that things tomorrow will be pretty much the same as things today, and that where they are not the operation of underlying similarities is responsible for any particular changes. We do not expect objects to gratuitously change shape, colour, size, and weight; we do not expect physical constants to suddenly vary, forces to spring up and die down. Our whole lives are premised on the stability of the natural world in myriads of respects. But now take a dimension in which, in given circumstances, we expect stability. We can then mechanically define a bent dimension of predicates governed by bent rules, with the feature that if things stay the same in our respect they change in the bent respect, and vice versa. For example, take the bent predicate red[BR] mentioned above; we would expect new rubies, mined after 1985, to be red; we would expect new wounds to give red blood. If so, they are not, like rubies and drops of blood observed before 1986, red[BR]. Since they issue after the crucial date, they would be red[BR] only if they were yellow. Now why do we prefer redness to red[BR]ness? What makes us think we have got hold of a similarity which nature

[2] *Fact, Fiction and Forecast*, ch. IV.

will herself protect, whereas some unfortunate who catches on to a bent rule will get a nasty shock in the dawn of 1986?[3]

Goodman himself, and many philosophers influenced by him, saw nothing but linguistic convention to hold onto. We do speak a language which sees redness as simple and red[BR]ness as bent; we could have spoken a language which placed them the other way round. But since we do and have used the former, we are in the habit of expecting things to continue red and of regarding any other prediction as irrational. We ourselves would look just as irrational from the perspective of a community which "entrenched" a predicate following the bent rule. It just so happens there are no such communities. This answer of course presupposes that there is some fact making it true that we mean the one thing and not the other by our terms. In other words, it brings us no closer to solving the problem of how it is possible to mean a determinate thing by a term, which loomed in the last section. It still leaves it possible that we form a wooden community, there being no truth that any of us mean one thing and not another by our predicates. Or perhaps we form what we call a mutually bent community, in which individuals have taken their training and the initial samples in different ways, formulating different rules which may suddenly dictate divergent applications, to our mutual consternation. But let us put that on one side. It is still true that our preference for our rules and not their bent counterparts is hard to regard as a product of purely conventional arrangements. For it is incoherent, in a way which I shall later elaborate (chapter 7), both to be confident that, say, future blood samples will be red, but also to regard that confidence as the outcome of an arrangement which we merely happen to have hit upon, out of a selection of equally attractive arrangements. If the opinion had such an insubstantial ancestry, so disconnected from the way things are, it had better not be trusted. In other words, if Goodman's were the ultimate answer, we would have no defence against total scepticism about whether the world is stable in any of the respects which we rely upon. Butter might be red tomorrow,

[3] There are sciences where this kind of question matters. Econometrics is bedevilled by the fact that many equations or models of what varies with what in an economy may fit initial segments of data, but diverge in their predictions of what will happen when things change.

and blood yellow; either might sing to us or cause golden eggs to come out of thin air.

A slightly more substantial answer would be that nature looks after our groupings, ensuring that our natural dimensions of classification are ones which are stable. Perhaps an evolutionary explanation works, so that time and selection weed out groups with a tendency not to latch onto the right groupings, not to "carve nature at the joints". This answer restores the coherence which seems lacking in a pure conventionalism: we might reasonably retain confidence in a prediction if we felt that a history of pressures which selectively favour the successful had brought it about that we have that confidence. But the answer has its problems: a group using red^{BR} would have done just as well as us until now. It is only in 1986 that they will get their come-uppance. The evolutionary story might actually predict that there would be a multiplicity of predictions at any time, since only the pressure of the *future* redness (or red^{BR}ness, as the case may be) of things defeats bent-rule-followers. Surely liberal nature would have grown some?

At this point we might begin to suspect the way we are looking at the issue. How credible is the possibility of these bent-rule-followers? Perhaps we can't ultimately make sense of the stories. For instance, consider again the deviant interpretation of the arithmetical rule: add 2 up to 186, then add 7. What else is true of the learner who takes this rule in the bent way? Does he think that he is *going on in the same way* when he adds seven bricks to a pile to which he was previously adding two at a time? What if he can carry two bricks but not seven? Does he not notice the difference: is he not aware of the shambles he gets into? Does the round^{BR} operator not notice that wheels "like" the ones he was initially shown do not roll, but meet flat planes sometimes at a point and sometimes along a line, and so on? Worse still, these people normally notice what we would consider to be similar deviations. If the foreman tries to load the bent bricklayer with seven bricks after he has added 184 bricks he protests: this is not going on in the same way. But two bricks later, it is. Again, the bent interpreter of 'red' is normally aware of the sudden emerging of a yellow member of a class whose members were hitherto red. It is only if the yellow arrival times its entrance for midnight 1985 that he regards everything as

"going on in the same way". I do not myself believe that we can really conceive of this kind of sensibility. Certainly we would in practice cast around for other interpretations of these people: perhaps they are blind to certain changes; perhaps they suffer from mysterious lapses of memory; perhaps they aren't really talking about numbers, shapes, and colours at all, but have other features in their minds, features of which we are not aware, and cannot define even in a bent way. In this respect, we do not inhabit the same world.

If this remains just a point about what we would do, it suggests cognitive imperialism. *We* would find it difficult to imagine someone who takes the bent rule to be straight; *we* would be mystified by his words, would be inclined to reinterpret him, and charge him with blindness to various differences of things. But perhaps that is just us, locked into our own capacities and "similarity spaces", with our own particular imaginations, natures, or conventions. If this is all that can be said, then from a more objective standpoint, the bent-rule-follower is not only possible, but there is no sense in which we are doing something properly, which he is doing badly. We couldn't regard ourselves as having locked onto the *real* similarities amongst things, or the right way to classify things, about which he is mistaken. There would only be correctness relative to a scheme of classification. This is the position in the theory of universals (that is, of rules governing predicates) we know as *nominalism*.

Can we really see these problems with the bent-rule-followers as the outcome of some contingent, parochial, fact about ourselves? This is not as simple as it looks. We can press the difficult questions. For example, the bricklayer responds very differently if the foreman loads him with seven bricks after he has added 184 bricks from how he responds after he has added 186. It follows that he must know how many bricks he has added. Equally if this bricklayer is watching a film about such an episode he doesn't know whether the man is going on in the same way unless he knows how many bricks he has already added. There seem to be three possibilities. Perhaps the bent-rule-follower knows these things in some mysterious, innate, way. But then he is not separated from us by mere conventions, or even by normal differences of receptivity: he is separated by

total mystery! The second possibility is that he knows these things normally: when he is loading bricks he keeps track of the number he has added, and if he loses track he doesn't know how to keep on the same way: coming in the middle of a film he cannot describe whether the colleague adds two bricks at a time, since he cannot tell when he passes 186. But that just shows that he is using this number to mark a difference: he needs to know whether the *different* thing which happens after it happens at the right point. This destroys the intended symmetry with us, that he should think of everything as going on the same way. We don't need to be aware of how many bricks have already gone to know whether the man continues to add two at a time. Finally, perhaps there are no signs of normal or abnormal knowledge. It is just that after 186 he struggles with seven bricks and shows *no* awareness that anything is different. To him it is as though everything is the same. But how can this be? This bent-rule-follower *fails* to perceive what we regard as differences, if they happen at the right point. But that's just failure. How does it transmute into the picture of someone who is genuinely making different judgements, following a different rule? It builds an image only of someone who forgets what he was doing at a particular time, or is unaware of much of what is going on: if we generalized this alternative over a good number of normal predicates (so that the man shows the same insensitivity to changes of shape, weight, colour, number, . . .) we simply end up with someone who does not know the world about him.

All these are things which *we* say! Well, they are of course. But before this rekindles the nominalist flame, we might reflect that since they are things which we say, we ought also to say that we can make no sense of the possibility of a sensibility, a way of perceiving and classifying things in the world, which naturally operates in terms of the bent dimensions. Bent predicates are important in the philosophy of language (and, of course, in the philosophy of induction, and for that matter in any branch of philosophy where similarities and differences are important, such as ethics or the philosophy of mind) because they offer a way of filling out the vision that our descriptions are in some way arbitrary, conventional, parochial to us. If the filling does not help the vision, then something else might in

principle do so, but it remains true that the introduction of bent predicates does not do the work which was expected of it.

Let us return to the bent learner. All this does not mean that he is impossible. It only means that if he caught on to the bent rules he would have to do *more* than simply go astray at 186 or the beginning of 1985, or wherever. He would need to show some awareness of the bend in the function he takes '+ 2' to express or in the class of things properly described as 'red' or 'round'. If he did catch on in the bent way he might, for instance, argue with a foreman who asks him to add bricks two at a time. He would point out how unreasonable this would be after he has added 186 bricks. Society's amazement at his difficulty would bring to light the odd way he took the explanations, and we would expect some simple account of what went wrong, and expect to have a reasonably clear remedy, explaining again what the foreman intends. What the learner cannot do is *both* take the initial samples and explanations in the bent way, *and* show no awareness of the bend, for by the above arguments we only deceive ourselves if we think we can make sense of that. It follows too that we can make no real sense of the possibility that *we*, now, might form a mutually bent community, so that having latched onto quite different rules of use of terms, we are poised to diverge in mutually unintelligible ways when presented with new things to describe. I can certainly eye my fellows askance in a number of ways. I can mistrust their judgements, their memories, their real grasp of any principle for applying terms. And on occasion someone may get hold of the wrong end of the stick; if he is brighter or less bright than others he may perceive similarities between things compared with which other differences are unimportant, or perceive differences where we see only similarity. This can lead to scientific reforms and advances, for although these mistakes and dissatisfactions may at a time be undetected, they are in principle detectable. They involve different dispositions and ways of talking and acting, in advance of different occasions of application of a term.[4] So since my companions show no undue interest

[4] This is a point on which I diverge from Saul Kripke. In his superb discussion of this (*Wittgenstein on Rules and Private Language*) Kripke denies that dispositions uniquely fix the question of whether a person is faithful to the right rule or a bent rule; in other words, a bent-rule-follower need not have different dispositions. For a discussion see my 'The Individual Strikes Back', *Synthese* (1984).

in 186, no undue fear lest next year's motor cars will have round wheels, no concern to lay in a stock of observed, yellow, bandages to match 1986 blood, I know them not to interpret the words in the bent ways defined. But are we entitled to talk of intended meanings at all? Does the wooden, no-rule picture, the third hypothesis, now come into its own?

3. Wooden Communities: Uses and Ways of Life

This is the central problem of the later world of Wittgenstein. Until paragraph *198 in the *Investigations* he develops what we have termed the negative point: 'any interpretation still hangs in the air along with what it interpreted, and cannot give it any support.' The existence of a rule governing the application of a term is not created by the existence of further substitutes for the term, for they pose the same problem. Wittgenstein then squares up to the threatening paradox, that the wooden picture is inescapable, and that nothing can create the existence of this rule. A man who calls a taxi a bus or a banana red would be no more wrong than us; just different. His training gives him one kind of disposition, ours gives us another. In neither case can the contents of our minds provide a rule determining whether what we are saying is correct or incorrect. Any course of action (i.e. any application or withholding of terms) can be made out to "accord" with what went before, as the example of the bent learner shows, and this just means that there is no right or wrong in our sayings, and therefore no judgements are made.[5] Wittgenstein's answer to this wooden picture is well known. He wants to connect the fact of a term being governed by a genuine rule, determining correctness and incorrectness of application, with its *use*, with a *custom*, a *technique*, a *practice*, with the fact that the word is embedded in the "language games" and "form of life" of a community of people. So how do practices give rise to meaning, when it looked impossible that anything would?

One suggestion is this. There is indeed nothing but the

[5] The example of a pocket calculator may help. Suppose such a machine with the oddity that although it adds 2 correctly to numbers before 186, when instructed to do so for numbers above that, it adds 7. It computes the bent function correctly, the straight one incorrectly. But in itself it just shows numerals, and there is no correctness or incorrectness about it. Only from some outside point of view is there a truth that it is functioning well or badly.

continuous flow of utterance, and the continuous play of dispositions, which seemed to allow nothing better than the no-rule view. But after all the dispositions include tendencies to correct, criticize, and adjust deviations. So in a *community* a deviant who calls a taxi a bus, or a banana red, is criticized. His behaviour will give rise to inopportune actions on the part of other people. Taking his utterance as they naturally would, they will go wrong. If the man appears normal in most respects then they say he spoke falsely; if enough of his classifications are out of step with the rest of the group's they find him unintelligible. (They may speculate about the possibility of bent-rules, suggesting that he takes his terms in some genuinely different way. But for the reasons I developed in the last section, they are unlikely to make much of that.) The point is that the *normative* aspect of meaning – the fact that some applications of terms are *incorrect*, and that the rules prescribe what kind of thing is *correctly* described as what – emerges from mutual pressures towards conformity.

On this view a community, in its language-using practice, is like an orchestra without either conductor or score, but with a tendency to turn on players whose notes are discordant with a democratic attempt at harmony. The negative point is supposed to prevent us from believing that each individual player has his own score – his own private instruction how to describe things, which gives him a standard for applying terms.[6] If we take this orchestral metaphor seriously it implies that the lone individual, considered quite apart from any surrounding group, could not mean anything by his terms. The wooden, no-rule view would be inevitable, for nothing exists to give the solitary speaker standards of correctness. Suppose he faces a square and describes it as round. This seems right to him, and what else is there to create a standard by which he has made a mistake? "Whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'."[7]

This is the heart of the famous anti-private language argu-

[6] Idealists who were sensitive to these considerations (including Kant and T. H. Green) tended to see the standard for description as laid down by the infinite or absolute mind in which we each to some extent participate – corresponding to the antecedent instruction of a composer. This theory is not attractive: what voice tells me how to take the instructions of this mind? Why should I listen to it? And how did it get off the hook?   [7] *Philosophical Investigations*, § 258.

ment in the later Wittgenstein, which I explore further in the next section. But the orchestral metaphor looks dangerously unconvincing if it is supposed to have such a strong consequence. Remember that the initial suggestion was that the existence of a technique or practice in which a term is embedded is the magic ingredient which creates a meaning for it. It needs to be argued that the technique or practice which is necessary to create meaning has to be *public*. Why shouldn't the solitary individual embed terms in instructions to himself: instructions which contrive to give him a technique, to cope in some way or another with his world, although there is no pressure from surrounding speakers creating standards of correct ways of taking those instructions? It is easy to go through the thought-experiment of coming across such an individual. A solitary individual growing up in perfect isolation – a born Robinson Crusoe – might give all the appearances of following rules, including linguistic rules, and of having a practice which embodies a distinction between correct and incorrect performance. Indeed we can imagine cases in which we simply have to say this. An example due to Michael Dummett is that of a born Crusoe who over the years evolves a technique for solving a Rubik's cube washed onto his island. There is no way of regularly doing that by chance. You have to follow rules. Perhaps to help himself he creates symbols reminding him of what to do at various points, and appeals to these on the way through the cube; with these symbols he can do it, and deprived of them he cannot. Clearly he has the *practice* or *technique* which entitled him to be regarded as meaning various determinate things by these symbols.

Wittgenstein's followers have tended to divide on the issue of whether his solution to the problem of meaning denies meaning for born Crusoes or not. The difficulty is that if the argument does deny it then, as I have just suggested, it seems unconvincing (consider what else he might do to show us that he really means things by his signs and symbols, or follows countless rules in his practices); on the other hand, if he is allowed to be a rule-follower, the orchestral metaphor ceases to embody any solution to the problem. A compromise is suggested by Saul Kripke. He points out that we might indeed think of Crusoe as following rules, but all that follows is that *if* we do so, we are

"taking him into our community and applying our criteria for rule following to him".[8] But it is not clear what this means, nor whether it gives the community any particular prominence in the creation of meaning. An orchestra coming across a solitary player, concentrating hard and making noises, might well say that *if* he were with them, he would be doing well or badly making these noises. But on the "democratic harmony" theory, they could not say that he *is* doing well or badly since in isolation there is nothing for him to do well or badly. The problem Crusoe poses is that he *does* have a practice (follows a tune) *regardless* of how we or anybody else think of him. Of course, Kripke is right that when we say this we apply our own criteria for rule-following to him; it is our judgement that he is following a rule. But this does not bring our community or any community far enough into the picture. It would be *our* judgement that an island has a tree on it. But whether an island has a tree on it is quite independent of how we or any community describe it, or even of whether any community exists to describe it. On the face of it, the situation is the same with the solitary intelligent Crusoe, in which case he has rules, meanings, standards for applying terms in his own solitary state, and with no reference to any community.

The problem with Crusoe shows that we must not fall into the common trap of simply equating practice with public practice, if the notion is to give us the heartland of meaning. It will need arguing that, contrary to appearance, the practice of isolated individuals cannot count. In any case, if the practice of an individual in isolation is not enough to create the fact that his words have meaning, how is the practice of a lot of us together to create the fact that *our* words have meaning? We talked earlier of the norms which arise from mutual pressures towards conformity in description. But how exactly does group conformity relate to understanding a predicate? In the orchestral analogy, a player knows whether he is wrong by listening for concordance with the group, and nothing else matters. But when I judge something to be red I am certainly not offering a shorthand for the more elaborate judgement: 'this is what most members of my group would call 'red'.' This can be no general solution to the problem of the meaning of predicates. For to

[8] Kripke, op. cit., p. 110.

make this judgement I need to understand the more compli-
cated predicate 'what most members of my group would call
'red'.' This predicate allows just as much for bent rules and no
rules. If we apply the equation once more we get: 'This is what
most members of my group would call 'what most members of
my group would call 'red''; and we are off on a regress. Any
predicate 'P' transforms into 'P*': 'what most members of my
group would call 'P' ': this transforms into 'P**', and so on.
There is no solution to our problem here. The more complex
predicates are not even synonymous with the bases from which
they derive, for it need not be true that most members of my
group call things which are X, 'X'. They may make systematic
mistakes. And in any event, it is no easier to conceive of the
fugitive fact, that we are genuinely guided by principle, when
we think of the more complex predicates, than when we con-
sider the simplest ones. So the orchestral analogy cannot be
taken this way.

The point also damages another analogy in which the
practice of a community gives rise to certain kinds of fact. This
is the analogy with the conventions which underlie *money*. Pieces
of paper can be of no value to an individual in isolation. But the
practice of a group of individuals can *create* the fact that their
pieces of paper have value to them individually. (Paper money
astonished Marco Polo when he visited China.) The analogy is
short-winded: the value to me of a banknote is *directly* a matter of
what other people will do for it. But in applying a word I am not
directly concerned with the reactions of other people. I do not
generally consider their assent or dissent to be the final court of
appeal on whether I am right. I am directly concerned with
whether a thing is red or round or whatever, which is a quite
separate issue from whether people describe it as such.

The democratic harmony view is responsible for the rela-
tivism which is frequently associated with the later work of
Wittgenstein (quite against his own intentions). An individual
player in the orchestra may go wrong. But how can the orchestra
itself do so? There seems to be no external standard by which it
can be deemed to be doing well or badly. The standard for
correctness in description seems to have shifted, as it were, from
conformity with how things are, to conformity with each other.
If I live in a community which calls the earth flat, is not this just

one more element of the dance, on which I ought not to get out
of step? There are actual illustrations of the way in which the
emphasis on practices and customs introduces this danger.
Suppose a group has a religion. Part of the religious practice will
be to say certain things – that the God or Gods are thus-and-so,
that various actions need doing, various doctrines are true. The
practice is to say these things; perhaps the practice stands the
group in good stead practically or emotionally. Saying the
things is an action: if it works, how can it be criticized? The
natural opposing thought is that if in these sayings they *intend to
describe* what the world is like, then they may be wrong. But
remember that it is the nature of the practice which is being
held to determine what the intentions are, not the other way
round. It is not their mental lives which determine the correct-
ness or incorrectness of saying that God is thus-and-so, but the
nature of their customs, techniques, ways of life. So they derive
the powers of these sayings from their role in their customs. And
then there seems to be no room for an ingredient of meaning
which makes it possible for the sayings to be *false*.

If this conclusion were right, it would best be taken to show
that the notion of a practice is an insufficient source of
standards of correctness, that is, of rule-following and of mean-
ing. If practices do not lift sayings into a *normative* dimension[9] in
which they are susceptible of falsity, and hence of truth too, they
are not filling the role which is demanded of them. The example
illustrates that we cannot glibly announce that the concept of a
custom or practice obviously has this power: it is going to be
difficult to picture the emergence of truth and falsity out of
customs and practices, just as it is difficult to picture the emerg-
ence of meaning out of any amalgam of mental and physical
facts. But the descent into relativism can be avoided. If the *only*
ingredient in the practice were to say the words, then it is indeed
hard to see why they should be taken as expressions of belief,
and susceptible of truth and falsity. And we have already
learned to doubt the authority of the person using the words: it
is not clear that he will have privileged access to whether they
express a belief, or serve some other role.[10] However the

* This normative aspects of things is stressed in Kripke.
[10] Suppose you say to yourself 'I believe in life after death'. Do you know that this
expresses a belief? Why not suppose that it expresses an attitude, or vague emotion?

practices may have many strands: the words used in religious ceremonies also occur elsewhere; the sayings are subject to the same kinds of criticisms and doubts as others, perhaps the commitments they express influence people in the same kind of way as other beliefs, and so on. The idea will be that a saying expresses a judgement if the practice involves taking it as expressing a judgement, which itself involves *procedures* of assessment, acceptance, and rejection. This, at least, ought to be Wittgenstein's answer. Ironically it follows that the major industry of taking his later work to release religious faiths from arguments concerning their likely falsity, is misdirected. The price of the release is that the sayings no longer express beliefs, for without the controls which are part of a genuine judgemental practice, language is just on holiday.

Crusoe showed that it is not clear what the word 'public' is doing if 'public practice' is regarded as the source of standards of application of terms. The religious example shows that it is not very obvious what counts as a practice either. In particular, we must be careful over what counts as *identity* of practice. In the last section I introduced the possibility of a mutually bent community, in which each individual had taken his initial exposure to terms in a different way. I urged, against nominalism, that although an individual could catch onto a bent rule, we could make no sense of his both doing this and failing to appreciate the bend. And this appreciation would display itself in different dispositions, or practices (the brick-layer who is asked to carry bricks up two at a time). Suppose, however, that I were wrong about this. In that case a *concealed* mutually bent community would appear to be a possibility, in which the apparent identity of practice at a time is a cover for different individual ways of taking terms, each supporting its own standard of future application. Each member of the orchestra would be resolutely following his own conception of how the theme should go, and the harmony in the first few bars would be a matter of luck. If a concealed mutually bent community is a possibility, then their common practice at a point seems to be no source at all for standards of future correctness. At a point of divergence the individual carries on one way, and others carry on other ways. There is just nothing to say who is "right", since their preceding practice allows for this diversity:

their tunes can be continued whichever way they see fit. The orchestra is in no position to criticize any individual player, since the democracy is no longer speaking with one voice.

Of course, we do not believe ourselves to form a concealed mutually bent community, partly because we believe in the common nature of mankind, and partly perhaps for the anti-nominalist reasons I developed in the last section. Wittgenstein might be seen too as denying the mere possibility of such a community. It gives each individual a conception of the right way the tune should go, or in other words his own previous intention to use a term in some specific, determinate way, and this intention exists and determines a standard for truth in his judgements entirely without reference to other people. This is in many commentators' eyes exactly the idea which Wittgenstein opposes, substituting instead either the democratic harmony view, or some close cousin which we come to later.[11] But it is still at this stage unclear why he can oppose it. Certainly the negative point warns us off one particular conception of this individual intention. If the individual has this determinate intention, and knows what it is, this is not made true by the presence to his mind of a particular display. But on the face of it that leaves other possibilities. Crusoe may know how he intends to use the symbolism which determines the way to solve Rubik's cube not just because particular pictures come into his mind, or other symbols, but because he does something which counts, for him, as according with the rule, and something which does not, if he makes a mistake. He has his own practice. Similarly we naturally think of the child of the last section, who has a rule for developing a series, as aware of how he *should* go on *by his own lights*. It is this determinate intention which gives him standards of correctness and incorrectness, and generates the truth that he means something, and is not, like a wooden child, merely writing numbers one after the other. This natural picture is not destroyed by the negative point. That only attacks one conception of what it is to have a determinate intention (it is to have some presence in the mind) and how we know of it when we have one (by introspective awareness). It is not by itself strong enough to suggest that no conception of the difference between

[11] See Crispin Wright, *Wittgenstein on the Foundations of Mathematics*, pp. 20 ff. for a good discussion of this.

the solitary understander, and the wooden individual, is possible.

The upshot is that it is no easy matter to make public practice or custom into the magic ingredient which would turn the wooden picture into the full one. The negative point offers nothing strong enough to oppose Crusoe's claim to be using and understanding a term, and to know its meaning through first-person knowledge of his own rules, intentions, and procedures. The negative point only shows that this knowledge is not to be thought of as a simple matter of the presence of some display in the mind.

My Crusoe would have invented a fragment of a personal language – an "idiolect". Now the fact that he is possible should not be taken to imply that a linguistic community may be regarded as a group of individuals, each with their own idiolect, but amongst whom, because of their need to communicate, there arises a pronounced similarity of idiolect. This conformity would be in a sense accidental: the fact making a word mean what it does in an individual mouth would be entirely a fact about the speaker, and a fact upon which he would be the authority. Whereas in an actual linguistic community we recognize independent authority. What a word means, and what a person has said by using words, is a socially fixed matter, and often does not accord with a speaker's own understanding, or lack of it. If a man says that he has an elm in his garden, or that his father has arthritis, he may have a very poor understanding of his own saying: he may understand no more than that he has some kind of tree in his garden, or that his father has some kind of ache in his joints. He may himself be quite unable to tell an elm from a beech, or arthritis from rheumatism: nevertheless we will not hold him to have spoken truly if he has a beech in his garden, or if his father has rheumatism. We enforce what in the next chapter I call deferential conventions, meaning that we recognize community authority, and expert authority, in providing the actual sense of words (4.6). We emphatically do not allow that someone has spoken truly because in his private idiolect 'elm' covers beeches as well. The reason why we do not allow this is that it threatens the social utility of language: we need social norms towards confirmity of usage if we are to rely upon the messages made with words. Otherwise we could not

reliably tell what is to be understood by the utterances of an individual.

Although this point is undoubtedly correct, it must not be overstated. It does not deny that an idiolect could be a self-standing language. It leaves it open whether an isolated individual might have the determinate intentions or procedures to afford a sense to his terms quite without any question of deferring to the authority of anyone else. For example, if the meaning of a term were thought of along verificationist lines, as the kind of procedure or experience which determines whether a term applies, then an individual might himself invent and fix such a procedure or such an experience (but see next section). It is just that such an individual can never rely upon there being a determinate meaning to a term when he does not himself know what it is. He thereby differs from members of linguistic groups, who defer to others, as in the elm and arthritis example. I am here dissenting from Dummett's treatment of this issue.[12] Dummett correctly takes the way we bind ourselves by deferential conventions to show that "there is no describing any individual's employment of his words without account being taken of his willingness to subordinate his use to that generally agreed as correct". This is true. But he continues: "That is, one cannot so much as explain what an idiolect is without invoking the notion of a language considered as a social phenomenon." This is not true, or, if it is it needs a different support. For it is not equivalent to the first claim: it implies that an individual cannot do for himself what a society can do together (provide meanings for terms), whereas the first claim says only that an individual will not actually have done that, but will be bound by social facts which he will recognize, or ought to recognize, and which can split his actual meaning from his own understanding. The stronger claim needs the idea that meaning *has* to be social, and this goes beyond saying that it actually *is* social.[13]

The essence of the matter is that there seems to be no impossibility in an individual creating and abiding by his own rule of use of a term. But if an individual cannot do this, the fact that he is surrounded by others seems a doubtful source of help. The

[12] 'The Social Character of Meaning', in *Truth and Other Enigmas*, pp. 424–5. Also 'What is a Theory of Meaning?', in Guttenplan ed., *Mind and Language*, p. 135.

[13] I discuss deferential conventions further in chapter 4.

fact that there are lots of individually wooden people, forming a public group, is not itself calculated to suddenly transform them into a non-wooden community, sharing genuine identity of concepts, with real rules of application.

Crusoe's idiolect is not a private language, in the sense which that phrase bears in Wittgenstein's most famous development of these thoughts about meaning: the anti-private language argument. Crusoe's practice, for instance with his signs which help to solve the Rubik's cube, is only accidentally private. If Man Friday arrives, there is no reason why he should not be taught the same procedures and rules. But the private language which Wittgenstein opposes is private in a stronger sense. Its terms are given their meanings by reference to private episodes, such as sensations. We must now consider Wittgenstein's arguments against such a language, and the place of those arguments in philosophy.

4. Privacy and Practice

So far in this chapter we have tried to find what makes it true that either (i) a person is using a word in one definite sense, or (ii) he is using it with a different, bent interpretation, or (iii) he is not using it with a sense at all, but is merely parading the term under the impression that he is doing so. *Some* fact must determine which of these is true. We have accepted the negative point, that it is not a display to the mind which does it. We have cast doubt upon the idea that the actual presence of a communicating group is essential: whatever fact it is which marks the difference, it seems possible that an individual should satisfy it by himself. In this way his earlier self can transmit information for his later self to profit from, just as different members of a group can. In effect we are left with a pair of suggestions (not necessarily exclusive): it is the existence of a *practice* or *technique* which makes the difference, or it is the existence of a determinate intention, known to the speaker in whichever way we know about our own intentions, which makes the difference.

As is often the case in philosophy, a good way of exploring these ideas is to see what they rule out. The anti-private language considerations aim at this conclusion: no language can contain a term whose meaning (or sense) is constituted by a

connection the term has with a private item, which lies, or lay, solely in the mind of the individual who understands the term. We can call the doctrine that rules out such a term, "semantic externalism".[14] It is a doctrine about the terms of any language at all, including ours. It tells us that no term of any language, including terms like 'pain' 'tickle' 'experience as of seeing red', have their meanings fixed by a certain kind of connection. The importance of this claim is that the reverse doctrine is so tempting. It is tempting to say that I know what a term like 'pain' or 'burnt taste' means from my own case. Under some circumstances I have a certain kind of experience. Others, supposing that I have it because of my situation or my reactions, teach me to use some word to apply to it. I absorb their teaching by giving myself a private ostensive definition: I focus upon the sensation, say a pain, and promise to call just that kind of sensation 'pain' in the future. It is the fact that I use the term in conformity to that rule which identifies its meaning. The rule fixes the connection between the term and the private sensation, lying solely in my mind. Semantic externalism opposes this model. By opposing it, the doctrine threatens whole clusters of ideas in the philosophy of knowledge and the philosophy of mind. It stands against the thought that our best or fundamental knowledge is of the contents of our own minds. It stands against the whole Lockean model of language, whereby the immediate significance of a word is an Idea in the mind of the person who apprehends it: what Locke believes to be true of all words, semantic externalism believes to be true of none. It eventually alters the whole conception of the privacy of our own experiences and sensations, although the consequences here are indirect, and need a little explanation.

Semantic externalism says that terms like 'pain' do not have meanings which are constituted by their connection with a private item. Now we might decide that pains, experiences, sensations, are precisely items of this proscribed private kind. In that case the doctrine forces us to revise the idea that it is through a connection with such things that any words have their meanings. Alternatively, we might suppose that it is quite certainly through that connection that the words have their meanings. What could be more certain than that the word

14 I borrow the term from E. J. Craig.

'pain' means what it does because of its connection with *pain*? In that case we are forced to revise our philosophy of mind. Pains, sensations, and experiences then cannot be items of the proscribed kind, and we must seek an account on which they are not items of private acquaintance, but are thought of in some other way. On the first alternative the private language argument, and the semantic externalism which is its conclusion, remain doctrines about meaning. We can think of minds in a traditional ("Cartesian") way as private repositories of experience, sensations, perhaps intentions, of which the subject has a privileged acquaintance. But our ideas of what gives the meaning to terms like 'pain', 'experience of seeing red', and so on, have to be altered. On the second alternative it is the conception of privacy itself which is threatened.

The first alternative seeks to reconcile Cartesian privacy with semantic externalism. Its strategy is to distinguish between the *sense* or meaning of a term like 'pain', which is not given by its connection with the private experience, and its *reference*, which may yet be the private content. But although the sense/reference distinction is quite legitimate (see chapter 9) it cannot effect this marriage. The distinction is at its most visible when we take phrases which can be fully understood (i.e. whose sense can be fully apprehended) when it is not known to whom or what they refer. Thus I perfectly understand many sentences containing definite descriptions ('the person who committed this crime'; 'the richest man in the world') although quite ignorant of who or what it is that they refer to (see 9.1). This is why I can understand and obey instructions like 'Look for the person who committed this crime', or statements like 'Economists do not know who is the richest man in the world'. The sense of these sentences is a function of the sense of the individual words occurring in the descriptions, and this is quite independent of whether *x* or *y* is the person who committed the crime, or the richest man in the world. But 'pain' does not function like this. There is no understanding of the term by people who do not *know what pain is*. It is through knowing what pain is that we come to understand pain ascriptions, whereas it is not through knowing who is the richest man in the world that we come to understand sentences using the description. Because of this we cannot separate out two processes: learning

the sense of the term 'pain', perhaps compatibly with semantic externalism, and *then* learning to what it refers (the Cartesian private item). Rather, it is by knowing what pain is that we come to understand the term, and what it means to apply it. If this knowledge is essentially knowledge of a privately shown item, lying solely in the mind of the subject, then semantic externalism is refuted.

For this reason, the second alternative is more promising: if we accept semantic externalism, we must revise our whole conception of the privacy of the mental, and our knowledge of the contents of our own minds. But what is there to be said for semantic externalism? What is the force of the anti-private language considerations?

Let us consider a proposed case of private ostensive definition. A man has a certain kind of sensation. This sensation has a "phenomenal quality" which is known to him alone: *he* is aware of it, just by having it. He can attend to it, like it or dislike it, relish it, and, let us suppose, christen it. By this christening he (purportedly) provides himself with an *intended rule*: in the future call only *this* kind of sensation, '*S*'. This rule would determine what is correct application of the term '*S*' and what is incorrect. A later sensation with its own definite phenomenal quality would be rightly called '*S*' if it falls within the intended range of the term, and wrongly called '*S*' if it does not, but the subject mistakenly takes it to do so – perhaps by forgetting the actual nature of the original example. To use one of Wittgenstein's metaphors, the intended rule provides a measure or yardstick to lay alongside a new sensation, which will then conform with it or not, as the case may be. Let us say that a man is *faithful* to the original christening if there was an original episode of ostensive definition of this kind, which gave him a definite intention to call only sensations of a certain kind of quality '*S*', and if the man later uses that intention as a rule which allows some sensations to be '*S*', and disallows others. Wittgenstein's endeavour is to show that there can be no such truth as this: no truth that a man is (or is not) being faithful to the original episode of this kind. The appearance must be a sham, and hence the idea that a man is later *judging* a new sensation to be '*S*', or not to be '*S*', is also a sham. For genuine judgement demands faithfulness to a pre-existing rule.

Otherwise the later occasion shows nothing but the man making a new *decision* (I'll call this 'S'/I'll not call this 'S'). Such a decision would not be responsible to anything that happened previously. Hence it would not be correct or incorrect, and it cannot be regarded as the making of a judgement. For judging is something which is essentially capable of being correct or incorrect. In this case, according to Wittgenstein, nothing previous created a standard whereby this can be so. Once again: "whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'."

Why can there be no truth that a man is being faithful to an intended rule, whose content was fixed by the first sensation? When the later case arises he might say, "Ah, here is an 'S' sensation again, and he has the *impression* of the term 'S' having a definite sense, so that this remark makes a judgement. But being under the *impression* that you are following a rule is not sufficient to be *truly* following a rule. We have already met the no-rule hypothesis, or possibility of a subject who thinks that he is following a rule when he applies or withholds some term, but who is like a lunatic covering pages with "sums", or like the man whom Wittgenstein considers in § 237 of the *Investigations*, who intently follows a line with a pair of compasses, with one leg on the line, and the other following at a distance, but at a distance which he constantly alters by opening and shutting the compasses as he draws the points along. This man may think he is tracing a path defined by the first line, but not be doing so. For since nothing would be a violation of this rule, the hypothesis that there is a rule is mere show. A rule must allow some procedures and disallow others.

In the public case the "wooden" individual, whose use of a term is not rule-governed, can be detected, because his practice is eventually different from that of someone making genuine judgement with the term. The lunatic's "sums" form no part of the practice of an applied mathematics (if they do, we might revise the opinion that there is no method in them). But in the private case, only the subject himself is an authority on whether his applications of the term conform to an intended rule. So Wittgenstein can ask what, in the private case, is the distinction between (a) someone who is genuinely faithful to a pre-established rule, which determines correct and incorrect appli-

cation of 'S', and (b) someone who is disposed to use the term under the illusion that he is following a rule determining its application?

At this point one is inclined to concentrate upon the phenomenology of the matter. Give yourself a sensation, remember it, and ask whether a later sensation is the same or different. It seems a well-formed, well-understood question, at least if you take care to specify various respects of sameness, like intensity, or felt location. But there is a possibility, even if it is one you are likely to dismiss, that you later misremember what the original sensation was, and hence misapprehend the intended rule it was used to introduce. You have then the new, candidate sensation, and a memory of the intended rule, fixed by the old exemplar. But the memory would be deceiving you. It would lead you to think that the new example is very like the old, and deserves the same name, when in fact it is quite different. Let us call this possibility (c).[15]

If Wittgenstein is allowed to use the verification principle, he is well placed to attack the idea that there is a real distinction betwen (a) and (b) and (c). For anything the subject does or experiences at a moment, or himself says, is compatible with each hypothesis. And the public is in no position to tell which is true either. No third person can tell whether the later sensations are really like the first, or really different, or whether the subject is really following no rule at all in what he calls 'S'. If there is no verifiable difference between the three hypotheses, then by the verification principle there is no real difference between them. But for the term 'S' to be meaningful there must be a difference between them, for it must be rule-governed and permit of incorrect application.

What is much more doubtful is whether Wittgenstein can reach this conclusion without relying upon a verificationist step. Many writers suppose he can.[16] They think that the challenge to say what *makes* the difference, in the private case, has its own force. It is not just a question of how we might tell

[15] "Always get rid of the idea of the private object in this way: assume that it constantly changes, but that you do not notice the change because your memory constantly deceives you". *Philosophical Investigations*, II, p. 207.

[16] e.g. A. Kenny, 'The Verification Principle and the Private Language Argument', in O. R. Jones (1971); C. Peacocke, 'Rule Following: the Nature of Wittgenstein's Arguments', in Holtzman and Leich (1981).

which one is true, but of whether we have any conception of what would *make* one of them true. The challenge is to state this, and it is alleged that the would-be private linguist has no answer.

The challenge needs quite delicate handling, or it threatens to destroy public language as well. That is, in so far as the individual has difficulty in meeting it, it is also possible that a group does. We have already seen how hard it is for a group to defend themselves as being genuine rule-followers, sharing an identity of concept, rather than a mere number of wooden individuals, or a mutually bent community holding quite divergent interpretations of shared predicates. But even if we waive this problem, perhaps because we draw on the idea of a practice, or technique with a term, which saves the public, the prospects for the challenge are not all that bright. This is because it is rather hazy whether the challenge, to show what makes the difference, can legitimately be met just by repeating the description of the three cases. Suppose the private linguist defends himself by saying: "We already know what makes the difference. In the one case there is a rule, and it is determinate whether an application conforms to it; in the second case there is merely illusion; in the third case there is a misremembering of which rule was established. If there is a challenge to verify these hypotheses, then unfortunately it cannot be met. But that is often the way with sceptical challenges, and does not disturb the genuineness of the distinction." The challenger will impatiently reply that this is not good enough: he wants to be shown what the distinctions *consist in*, in the private case. But what does this mean? Perhaps only that the distinction should be drawn in other ways, themselves making no mention of intentions, or rules, or fidelity to a pre-established sense. But why should this request be legitimate? A distinction made with one kind of vocabulary often cannot be captured except by using that vocabulary: the distinction between red and green is essentially a distinction of colours, and cannot be shown to "consist in" some difference which does not refer to colours. The distinction between happiness and pleasure is a psychological distinction and cannot be made except in terms from that theoretical vocabulary, and so on. Notice too that by urging the negative point, Wittgenstein has already led us to think of intentions as

irreducible, in the sense that the issue of the intentions with which a person uses a term is never just the issue of which display he has before his mind.

So Wittgenstein's challenge is objectionable if it presupposes that the distinctions between (a), (b), and (c) must be capable of being drawn in other terms. It is also objectionable if it threatens public language as much as private language. Following Kripke we can put Wittgenstein's challenge in the garb of scepticism: this sceptic denies that anyone can *know* whether, after an attempted private ostensive definition, and attempted further use of the term, it is really (a), or (b), or (c) which is in force.[17] But then we have already discovered the scope for equivalent scepticism in the public case: the sceptic who wins against the private linguist looks well set to win against a public group when, corresponding to the three hypotheses in the private case, he asks whether the right-rule view, or the no-rule view, or the bent-rule view, is the true one. If the moral of the rule-following considerations is that we, the public, cannot meet this challenge except by insisting that we do know what we mean, and that we mean the same, and know this by knowing our intentions, then the would-be private linguist can avail himself of the same liberty.

If the would-be private linguist sits tight on his claim to have a determinate intention in calling a new sensation 'S', the only way forward is to concentrate upon the notion of a practice, or technique. Suppose we accept that understanding a term is possessing a skill with it, and that this skill is to be thought of as a kind of technique or ability. All terms of a language must be associated with such a technique. Then perhaps there is argument to show that the term 'S', introduced by private ostension, equips its user with no genuine technique at all. It forms no part of a practice whose proper pursuit stands the user in good stead, and whose improper pursuit leads to errors and disappoint-

---

[17] One must be careful of framing these issues around the figure of a sceptic. Wittgenstein's point is never to arrive at a conclusion of the form 'so we don't know whether . . .' His aim is to alter our conception of the facts we take ourselves to know: the aim is metaphysical. But his means to such conclusions may use sceptical dialogues as an integral part: *if* this conception of the facts were the right one, *then* we wouldn't know such and such, but we do, so we need this other conception of the facts. Again, I discuss this further in 'The Individual Strikes Back'.

ments. At first sight this idea seems promising. For precisely because there is no verifiable difference between (a) and (b) and (c), it seems that a "mistake" in applying 'S' is utterly inert. Let us adopt the standpoint which Wittgenstein is attacking for a moment. Imagine two prospective private linguists, each giving themselves what is in fact the same ostension, and coming away with the same intention. One has a better memory or better luck than the other, and only applies 'S' thereafter to cases which do fall within the range of the original intention. The other is fickle and faithless and often applies 'S' to cases which should have been excluded. How is it that one does well and the other does badly? What cost does the errant linguist incur? Apparently, none whatsoever. And if this is so it suggests that there is no real failure of a technique or practice here. A technique is essentially something which has consequences, and whose failure can let us down. So the hypothesis that the would-be private linguist is really operating a technique seems to be pure show. The normal surroundings and stage-setting of the technique of judgement are missing: he is like a man driving an imaginary motor car, or playing an imaginary piano. Consequently it cannot *matter* whether we regard him as a case of (a) or (b) or (c).

The obstacle to this range of thought is that it is we who introduced the S-classification as an isolated and inert incident, with no consequences for good or ill. But suppose on the contrary that the private linguist's performance *is* part of a technique which he is forming, testing, trying to render reliable. The technique is to bring order into his life. By correlating the recurrence of one experience with the recurrence of others of related kinds (warmth – pain! visual experience x – tactile experience y, etc.), the private linguist can begin to find order in his subjective world, and an ordered subjective world is a nice thing to have, since without it we can have no understanding of ourselves as conscious of an objective or spatially extended world. So the enterprise to which the classification of private experience belongs can be a serious one, and it can go well or badly, and it can matter to the subject why it is going badly. If, for instance, S which is usually followed by R is on some occasion not so followed, the private linguist may doubt their general correlation, or doubt either or both of his classifications. And it can be a serious matter to decide where the failure lies.

and to accommodate future classifications and expectations around it.

In this circumstance there is no reason at all for the private linguist to take the attitude that whatever seems right is right. He may do better to take the attitude that his memory is not totally reliable, that it is easy to fail to notice genuine differences between 'S' and sensations like 'S' but importantly different in what surrounds them, and so on. Judgements of recurrence take their place as corrigible in the light of subsequent experience. The moral will be that there is a point to discriminating the private linguist's performance as genuine judgement, capable of truth or falsity in the light of pre-established intentions, only when the performance is part of some general technique of belief-formation. It would then follow that beliefs have to come in populations, but not that believers do. The would-be private linguist's title to think of himself as a believer would be derived from his title to think of himself as a theorist, attempting a whole set of views about the order of his mental life.

A philosopher impressed by the challenge to make the distinction between (a) and (b) and (c) will still complain. Perhaps we can suggest why the private linguist may pointfully take the *attitude* to himself, that he is making genuine judgement. He need not take the attitude that anything that seems right will be right, nor that his dispositions to call things 'S' answer to no previous intentions. They have the crucial normative dimensions of correctness or incorrectness. But, the opponent will ask, how is this attitude justified? If there is no *fact* of correct or incorrect applications of the term 'S' to a new sensation, then surely the attitude itself involves a delusion. It is not enough to say that a private linguist (or public group) may dignify or compliment himself on being a rule-follower or on making judgements which have a genuine dimension of correctness: the compliment must not be empty. There must be a fact of the matter whether he is one or not. And we still have no conception of how the original episode reaches out, as it were, to constrain the proper use of 'S' on any subsequent occasion.

I do not think that it is true to the later Wittgenstein to pursue the challenge this way.[18] Firstly, Wittgenstein is on the side of those philosophers who query the borderline between genuine

[18] This is another point of difference from Kripke.

beliefs (in *facts*, with *truth-conditions*) and other kinds of commitments (such as possession of attitudes, or acceptance of rules). The error of his earlier philosophy was to exalt a simple, single conception of a fact (as in effect a spatial array of objects), and to make no room for truths expressed in terms which do not refer to such spatial arrays: for instance, truths about causal relations, about psychology, about the will, or about ethics. The later work is acutely conscious of the way in which our difficulty in conceiving of psychological facts arises from a spatial or physical model of what a fact must consist in. This is why we find our ability to think of absent things or to form intentions which cover cases which we have not thought of, so mysterious. But the characteristic tone of the later work is one of toleration towards different vocabularies, even when we "have no model" of the truths they describe.[19] A pertinent example is Wittgenstein's famous reaction to the talk of the belief that another person is conscious: "My attitude towards him is an attitude towards a soul. I am not of the *opinion* that he has a soul."[20] The belief or attitude is not a belief *in* a certain fact (a spiritual interior to the animal, or a ghost in the machine), but is something whose content is given by my reactions to the person, ways of behaving and dealing with him. Against this background it would not be appropriate for Wittgenstein to insist that the private linguist can *only* take up an attitude towards his own classifications (the attitude that they answer to a previous rule or intention), but that there can be no *fact* about whether they do: the justifiable attitude is just the kind of thing to give content to this fact. (I offer my own exploration of the attitude/belief distinction in chapter 6.)

When the would-be private linguist classifies a new sensation as '*S*', he thinks of himself as laying it alongside his rule. The rule came into his possession after the private ostension; its content is given in his intention which was formed on that occasion, to call only things like the original example by the term. Wittgenstein's brilliant strategy was to shrink this alleged

---

[19] *Philosophical Investigations*, §192. Cf also the discussion in *Remarks on the Philosophy of Psychology*, vol. 1, when Wittgenstein says of expressions of intention: "Yes, and such use of language is remarkable, peculiar, when one is adjusted only to consider the description of physical objects" (§ 1137).

[20] *Philosophical Investigations*, p. 178.

---

comparison down to a point, leaving nothing but the new sensation, and a bare disposition to say '*S*' or not. The previous history drops out from the use of the term, for everything about that use could be the same whether or not your memory is utterly deceiving you about the character of the original example (case (c)), or whether you are only under the impression that you are really guided by the nature of the original example (case (b)). The difficulty is to destroy our stubborn conviction that we are right to believe ourselves to be in one category and not in either of the others. Broadly speaking there are two possibilities. Wittgenstein can put on the garb of a sceptic, and allege that we do not know the nature of the previous episode, and the intention which it was used to form. Or, he can raise the metaphysical or ontological charge that we have no conception of the fact that we are genuinely guided by the previously formed intention. Either way, the issue is likely to remain inconclusive. The sceptical charge is too near to blockbuster scepticism, which would destroy our knowledge of anything, especially of public meanings. The ontological or metaphysical charge is too near to insisting that there should be a fact of our being guided by a previous intention which can, as it were, be laid out to view in the form of an image in the head or other guide. This is just what the negative point attacks. And then, if the private linguist refuses to try to force his fact into this kind of shape, it is hard for the later Wittgenstein to deny his right to do it.

5. *Exercising Mental Concepts*

Even if the anti-private language argument is inconclusive, it does a tremendous service in the theory of knowledge. It entirely subverts the idea that our knowledge of our own meanings, derived from the acquaintances we have with our own mental lives, is a privileged, immediate, knowledge, beyond which lies only sceptic-ridden insecurity. This Cartesian picture, according to which my knowledge that my present sensation is green, or hunger, or a headache, is peculiarly incorrigible, is, I believe, overthrown by the realization that the only "incorrigible" element is the single point, the present sensation; any enterprise of judging it to be one thing or another

involves bringing it into contact with a rule or previous intention, and hence brings the possibility of misidentification of what that was. But this does not entail that there could be no such judgement. It only entails that it has no unique immunity from error.

There is one further aspect of this difficult area which needs mentioning. I characterized semantic externalism as the denial that the meaning of any term can be "constituted by its connection" with a private exemplar. The same thought is sometimes put by denying that it is "from our own case" that we learn the meaning of sensation terms. The alternative account of the meaning of such terms is likely to stress public criteria of their applicability. To understand a sensation term requires knowing what kind of situation or what kind of display would make it appropriate to attribute the sensation to a third person: it might also demand knowing that people can sham, or be hypnotized, or whatever, into appearing to have sensations when they do not. But this, it is suggested, gives no ground to Cartesian privacy. It only shows that our public evidence is inconclusive or "defeasible" by further public evidence, which could arise if peoples' motives for shamming, or queer state of hypnosis, were removed.

Accuracy demands a little suspicion of some of this. There is at least room for a theory which admits that it is not *wholly* from one's own case that one understands a sensation term, or that the meaning of such a term is not *wholly* a question of its connection with the private example, but also insists that it is *partly* from one's own case that one understands the term. That is, a full grasp of sensation terms would require knowing what the sensation is like (privately), and realizing that other subjects might or sometimes do possess it too. It might also require understanding the kind of evidence which justifies the belief that they do. It is an awkwardness of the anti-private language argument that it seems to rule out even this diluted semantic internalism. It is so strong that it leaves *no* role for the internal examplar. For the subject can only be under the impression that his own example of the sensation plays any role at all in identifying a rule of application, and this impression is, supposedly, not enough to make it true that it plays a role. And nothing else can make that true either.

Even dilute semantic internalism faces or raises difficulties. So far I have concentrated upon what we can call a "vertical" version of the private language argument. This considers one agent and his relation to his past states. But there is a "horizontal" aspect of the argument too. This questions whether dilute semantic internalism can allow any sense to the thought that some other subject has the same kind of sensation as oneself. To make the problem vivid, imagine someone arguing that he does indeed get the concept of pain from his own case; his own case enables him to tell when there is pain about, which is to say, when *his* body is injured or he is affected in some unfortunate way; hence no other events are associated with pain at all. *Your* injury is just not the sort of thing which causes this sensation. *He* never feels it when *you* are injured. The concept of sensation is exercised "vertically" as it were: in the one dimension of my own feelings.

The challenge is to explain how, *if* our basic use of the concept is in our own case, we can ever come to exercise it in full generality. How do we understand that we are each just one of the many creatures which equally have sensations (even if we then go on to wonder how like our own those of our fellows might be)? This kind of challenge is of great importance in philosophy. For example, Berkeley anticipated Hume in finding it hard to understand how we gained a concept of causation by acquaintance with the ("passive") flow of events in the world which we sense. He proposed instead that the origin of the idea lies in our knowledge of our own agency or exercises of will.[21] But (as he realized) that makes it impossible to see how we can properly describe non-mental things as causing anything. Again Hume describes the origin of our idea of justice in the need for a scheme or system of rules whereby we can gain reciprocal advantages from one another: there is a problem then whether he can make sense of our idea that we have a duty of justice to future generations, or animals, who cannot reciprocate.[22] We could put the challenge by asking why, on these theories, the term in question is not ambiguous, meaning one thing in the home case (from which we get the idea) and another thing in the further cases. How can it be the *same*

---

[21] *Principles of Human Knowledge*, § 25.
[22] *Treatise of Human Nature*, Bk. III, Pt. II, Sect. 1.

concept, which is exercised in the home case and in the further cases?

If the challenge were just one of explaining the genesis of our understanding, it would seem feeble. We might reply that it just comes naturally to us to extend the concept from the one kind of case to the other. But the real challenge is to say why it is the same concept in each case. Is it responsive to the same kind of evidence or argument; do we appreciate its consequences in the same way? In the current example a dilute semantic internalist cannot just say, for instance, that in my case I recognize and remember my private exemplars, but in your case I exercise the concept of pain by, for example, caring about you or reacting with emotion when I judge that you are in pain.[23] He must go on to *connect* these two exercises: to show why it is appropriate to talk of *one* concept which I apply equally to you and to me. However there are steps towards meeting this demand which the semantic internalist can take. He can point out that many of the things I know (or believe) about my own pains I also believe about yours: that it is no accident that this kind of sensation makes me behave in this kind of way; that there are ways in which I can conceal sensations, that sometimes I cannot, that there are ways for you to behave towards me if you believe me to have a sensation and also have various attitudes towards me, and that these are ways I behave towards you in the light of the same beliefs. In other words, I can judge so that my sensations and yours are subjects of similar predicates and claim that the same concept is involved just because of that. It would, I believe, be extremely hard to phrase the demand or challenge so that this kind of answer does not meet it. Consider this parallel: people think it makes sense to ask whether numbers are *objects*, but how can the question even make sense when the notion of an object is at home in talking of ordinary spatio-temporally located and bounded, solid and visible things? Answer: because (perhaps) we say enough about both numbers and more ordinary objects to explain the common term (they are equally referred to, counted, known about, independent of us . . .)[24]

So semantic internalism can attach a meaning to the view

[23] Although curiously this is at least a part of Wittgenstein's own thought.

[24] So we do have identity of meaning, or only similies or metaphors? Numbers are like objects in that . . .; unlike them in that . . . See more on this in 5.7.

that other people have sensations like ours. This leaves the question of whether we know that they do, and how we rebut scepticism about their similarity to us. But that takes us too far from the philosophy of language. The present verdict is that the private-language considerations seem at best inconclusive. There is no compelling reason why there cannot be a practice of judging that our own private sensations are thus-and-so. And the intention with which we apply the classification may, so far as the argument goes, be identified by private ostension. There is equally no compelling reason why such a practice should not also serve to identify (part of) the meaning of our public sensation terms. The two elements of meaning which we have been forced to make prominent in coming to this conclusion are, firstly, the relation between meaning and intention, and secondly the relation between meaning and a whole practice of coping with the world. I propose to pursue the first of these in the next chapter, and the second infuses the next part of the book, in which we consider ways in which we judge the world, and the kinds of truth they deliver.

Notes to Chapter 3

3.1   The central references for this chapter are:

L. Wittgenstein, *Philosophical Investigations* §§134–230.
N. Goodman, *Fact, Fiction and Forecast*, ch. IV.
B. Russell, *Human Knowledge, its Scope and Limits*, pp. 422 ff.

Russell makes only passing reference to bent rules. Goodman uses them to shed new light on the problem of induction, but bent rules and no rules are used to cast their shadow over meaning by Wittgenstein. Easily the best commentary to date is Saul Kripke's, referred to in the text (n. 4 above).

"It does not follow that the subject himself is *not* an authority . . ." It is vital to see that Wittgenstein is not denying that there is such a phenomenon as rule-following (nor that there are facts about which rule is in force, not that it is true or false that there are rules in force). The whole debate is about the conception of these facts that we can obtain. Even the best modern commentators (including Wright and Kripke) do not bring this point out fully enough; they leave it uncertain whether we can properly allow facts of this kind. They have a good excuse, because we have here a classic philosophical problem: a

critique of one conception of a kind of fact is so powerful that it leaves people unsettled whether we can any longer go on saying the things we used to say about the area. This predicament is discussed further in 5.1 and 5.2.

3.2   "A slightly more substantial answer . . ."
Although I say enough in the text to register my disagreement with both conventionalist and naturalistic approaches to Goodman's problem, a fuller discussion would bring in much more. I discuss the paradox with reference to induction in *Reason and Prediction*, ch. 4. Important discussions of Goodman's paradox include:

S. Barker and P. Achinstein, 'On the New Riddle of Induction', *Philosophical Review* (1960).
J. J. Thomson, 'Grue', *Journal of Philosophy* (1966).
P. Teller, 'Goodman's Theory of Projection', *British Journal for the Philosophy of Science* (1969).
S. Shoemaker, 'On Projecting the Unprojectible', *Philosophical Review* (1975).

See also 7.7

3.3   "A compromise is suggested by Saul Kripke . . ."
I give a more detailed discussion of Kripke's views in 'The Individual Strikes Back', *Synthese* (1984) (this volume, edited by Wright, also contains relevant papers by J. McDowell and C. Wright).

". . . taking his later work to release religious faiths . . ."
There is a trenchant discussion of post-Wittgensteinian views of religion in J. Mackie, *The Miracle of Theism*, ch. 12. It is, in my view, very uncertain whether the stress of religion as a human phenomenon, giving vent to human needs and feelings, actually conflicts with the view that religious beliefs are real beliefs, capable of truth or (usually) falsity.

3.4   Any selection out of the huge exegetical and critical literature on the private language argument is bound to be fairly arbitrary. Treatments which should profit students, in addition to those already mentioned, include:

A. J. Ayer, 'Could Language be Invented by a Robinson Crusoe', in O. R. Jones (1971).
R. Fogelin, *Wittgenstein* (The Arguments of the Philosophers), chs. XII and XIII.
J. J. Thomson, 'The Verification Principle and the Private Language Argument', in O. R. Jones (1971).

C. Peacocke, 'Rule-Following: The Nature of Wittgenstein's Arguments', in Holtzman and Leich (1981).

But the best modern discussions, triggered by Kripke, seem to me to be nearer to the real heart of Wittgenstein's problems.

"I do not think it is true to the later . . ."
I here agree with two other writers who have explored Wittgenstein's later conception of fact:

P. Winch, 'Im Anfang war die Tat', and
B. McGuiness, 'The So-Called Realism of Wittgenstein's Tractatus', both in Block, ed. (1981).

3.5   "horizontal" and "vertical"
This adverts to the extensive literature which believes Wittgenstein to have subverted the familiar argument from analogy for other minds. The idea is that we cannot learn what pain is, in our own case, and then as much as *understand* what it could be for someone else to be in pain; hence we cannot argue by analogy with our own case that other, similarly behaving bodies belong to subjects with similar mental experiences: we cannot understand what this means. A good collection on this problem is *The Philosophy of Mind*, ed. V. C. Chappell, especially the editor's introduction, and the papers by Malcolm and Strawson.

I am very conscious that this section suggests more profound problems than it manages to treat. It is probably fair to say that the philosophical community, at present, is involved in a general shift away from supposing that anything is ever learned in our own case – this being supposed to involve a "Cartesian" conception of mind, whereby the contents of our own minds are immediately present, private showings, whose nature wholly determines our thought. The whole difficulty is to separate what is right about the direction, from what is wrong, or questionable, about the individual theses which people have taken to support the direction. Although I am sceptical of some of the arguments from the philosophy of language which have been used to support the shift, I have no settled opinion on the shift itself (the time will come when people begin to ask what was *right* about Cartesian intuitions in the philosophy of mind). In this work I am only concerned with the arguments as they emerge in connection with understanding and meaning. This is also evident in chapter 9, where in order to avoid a Cartesian, or even solipsistic view about the nature of thought, philosophers have been led to defend unduly implausible views about reference and thought.