CHAPTER 8

DOES *PHILOSOPHICAL INVESTIGATIONS* I.258–60 SUGGEST A COGENT ARGUMENT AGAINST PRIVATE LANGUAGE?*

CRISPIN WRIGHT

1. *What is a 'Private' Language?*

As the notion is usually understood in the literature, a private language is one which, necessarily, only one person can understand. Wittgenstein's own remarks (I. 243, 256)[1] may encourage such an interpretation. But it is not quite right. Intuitively, two people share an understanding of a predicate if what qualifies an item to fall within its extension is the same for both of them. Accordingly, if I somehow invented a language apt for the description of material—sensations, or whatever—in principle accessible only to myself, someone else might nevertheless understand the language: he would do so if he associated with its various descriptions material of the same respective kinds as I associated with them. This is just the natural and familiar thought that, while the phenomenological qualia associated by each of us with the word 'pain' are unknowable by anyone else, and are regarded as playing a constitutive role in our

 This material originated in one section of a paper on Kripke's interpretation presented to the 7th International Wittgenstein Symposium at Kirchberg-am-Wechsel, Austria, in 1982, and grew under the stimulus of seminars given in 1983–4 and 1984–5 at the Universities of Pennsylvania, Glasgow, Cambridge, Durham, Lampeter, Leeds, and Edinburgh; and at a Wittgenstein 'Workshop' held in Paris under the auspices of the Collège International de Philosophie. I am grateful to those who participated, for many helpful comments and criticisms, especially R. Rockingham Gill, Hugh Mellor, James Ross, Flint Schier, Bob Sharpe, John Skorupski, and Scott Weinstein. I should particularly like to thank Warren Goldfarb for his work, at my request, on the problems of the Appendix. The proof of claim (d) is due, in all essentials, to him. Finally, let me thank Simon Blackburn, Peter Carruthers, O. R. Jones, and Leslie Stevenson for useful comments on an earlier draft.

[1] All references are to *Philosophical Investigations* unless otherwise stated.

respective understandings of 'pain', there may nevertheless *be* mutual understanding if, as it happens, our qualia are appropriately similar. Only, if at all, in the case of proper names for which knowledge of reference is held necessary for understanding would the inaccessibility of my material to others be an essential bar to shared grasp.

A successful polemic against private language is meant, *inter alia*, to refute this picture of the meaning of sensation vocabulary. Since the picture does not entail that each of our idiolects of sensation is private in the standard sense,[2] we need to qualify somewhat the usual terms of the discussion. Private language had better be, not a language which necessarily only one person can understand, but a language which, necessarily, no two people can have adequate reason to believe they share. If your pain-quale is inaccessible to me and constitutive of your understanding of 'pain', what (uncontroversially) follows is not that your understanding cannot coincide with mine but only that I cannot have adequate reason to think that it does.

It merits remark that the inaccessibility, to others, of a speaker's material would not be the only possible source of private language (in the qualified sense). For the condition presented above as sufficient for shared understanding is also necessary: even if we are speaking of material accessible to everybody, shared understanding of an expression requires, in addition, that our uses of it be informed by the same conception of what qualifies an item to fall within its extension. So evidence of shared understanding has to involve evidence of a certain community of intention: evidence that we intend to let our respective uses of the expression be answerable to the same constraints. Now, a familiar sceptical line of reasoning aims to generate doubt whether a third party can ever have satisfactory evidence for the content of another's intention. If that were so, it would follow that intentions are, by their very nature, things which no two people can reasonably believe they share; so *all* language would be private, whether concerned with a publicly accessible subject matter or not.

The scepticism in question trades on the apparent first/third-person asymmetry in our knowledge of intention. *My* intentions (to

[2] As Edward Craig notes ('Privacy and Rule Following', in Jeremy Butterfield (ed.), *Language, Mind and Logic*, (CUP, Cambridge, 1986), this blocks any interpretation of Wittgenstein's argument which would build directly upon the impossibility of *consensus* in the description of private material.

prescind from consideration of the unconscious) are immediately, that is non-inferentially, available to me, and my beliefs about them are authoritative. For *your* intentions, on the other hand, I have only the evidence of what you say and do. What you say, however, is of no evidential use to me until it is interpreted; that is, until I have arrived at grounded hypotheses concerning the intentions which inform (your use of the ingredients in) your utterances. So your linguistic behaviour, it appears, can justify me in attributing a certain intention to you only if I assume knowledge of others of your intentions. Presumably, therefore, the *ultimate* basis for my beliefs about your intentions must lie in your non-linguistic and uninterpreted linguistic behaviour. And now it appears that this 'basis' must inevitably be hopelessly crude: that no end of alternative construals of your intentions, beliefs, and desires will be reconcilable with it. So, in outline, runs the sceptical argument: the prototype of the refinements respectively developed by Quine, in his writings on meaning and translation, and by Kripke, in his exegesis of Wittgenstein.

2. *The Significance of the Issue*

The philosophical consequences of the impossibility of private language will remain profound. First, solipsism will be an untenable position. Solipsism holds that no subject can have adequate reason to believe in the existence of any consciousness besides his own. It follows that no subject can have adequate reason to believe that he shares a language with another—since that would require belief in the *existence* of the other. From the standpoint of solipsism, language is essentially private; if there can be no such language, the solipsist is deprived of the medium in which to conduct his solipsistic dialectic. A demonstration of the impossibility of private language will therefore be a demonstration that there is error in any philosophy of mind, or epistemology, which has the consequence that the existence of another consciousness is at best a groundless assumption.

Second, a remodelling will be called for of the natural pre-philosophical conception, touched on above, of the kind of meaning possessed by our talk of sensations and other mental states. We tend to view the understanding each of us has of a word like 'pain' as possessing both a public and a subjective component. The public component is conceived as graspable by one incapable of feeling pain: it

is constituted by our shared concept of what pain-behaviour is and of the consequences, personal and social, of someone's being in pain. The subjective component, in contrast, is fixed by the character of painful experience; only one who can suffer pain can imbue his understanding of the word with such a component, and the component is, in the nature of the case, idiosyncratic. Now, the niche here granted to a public component may be held to obviate any implication that the language of sensation, and of the passions generally, is already, for each of us, a private language. But if the felt quality of my experience has *some* part to play in determining the content of the relevant parts of my vocabulary, and if it is accepted that this quality can be known only by myself, it must follow, it seems, that we cannot have reason to think that we *fully* understand each other's talk of sensations, and so on. Besides, if it is coherent to grant that material in principle accessible only to me contributes at all towards determining what I mean, what obstacle can there be to the fiction of a language in which it makes the *sole* contribution—a full-fledged private language in just the sense we are concerned with? The pre-philosophical conception, even if it escapes the outcome that the language of sensations is already, for each of us, private, must at least, it appears, be committed to the *possibility* of a private language. Accordingly, a demonstration of its impossibility will be a demonstration that this conception cannot contain the germ of a satisfactory philosophical understanding of the language of mind.

The issue is potentially of significance in at least one other important respect, connected with the sceptical argument adumbrated above. The first/third-person asymmetry which we are tempted to find in the epistemology of intention is no less tempting, of course, in the case of a speaker's meaning; what *I* mean by an expression is available to me directly and with special surety, whereas the meaning *you* assign to that expression is, for me, a theoretical construct, earned by inference, and as precarious as any (strictly) unverifiable hypothesis. Once speaker's meaning is so conceived, we can hardly avoid thinking of a communal language as itself a theoretical postulate; something which, if it exists at all, is constituted by the overlapping of first-person transparent idiolects, but whose existence can be at best a good conjecture. Now, if the sceptical argument succeeds, 'good conjecture' has to be replaced by 'assumption'; and communal languages, if any exist, will have to be constituted by overlappings of private languages. So if the sort of sceptical argument adumbrated is

sound or, more generally, if the first/third-person asymmetrical view destabilizes, under one form of pressure or another, to a point where only unjustifiable conjecture is available to a third person, then a demonstration of the impossibility of private language will accomplish a proof that the intuitive asymmetry is already a misconception; that no priority can coherently be accorded to the notion of an idiolect, conceived as first-person transparent, in an account of what a *communal* language is.[3]

The issue, then, really is a pivotal one in general epistemology, in the philosophy of mind, and in the philosophy of language. Not that I think that many are under any illusion about that. But a reminder may be salutary, when so much of the literature concentrates on its success or failure, of how deep the roots of significance of Wittgenstein's argument penetrate. And we need to be clear, as the above, I hope, makes it clear, that this depth is in no way compromised by the marginal re-interpretation of 'private language' with which I am going to work.

3. *Constraints on a Cogent Argument against Private Language*

A genuinely cogent argument against private language will have to observe a number of special constraints.

First, whatever the detail of the argument, it had better not have the additional strength sufficient to make trouble for Robinson Crusoe. A private language is to be a language which no two people can *in principle* have adequate reason to believe they share: if the argument seems to show that merely *practical* obstacles—like desert island isolation—defeat the possibility of language, then, even if we can find no fault in it, it will fail of cogency because seeming to prove too much. A cogent argument should be effective in the moulding of reasonable people's opinions. So no argument which generates a *paradox*, however seemingly watertight, can be cogent in the sense we require. To suppose that a solitary individual, even if isolated from birth, could not invent and utilize consistently (or as nearly consistently as need be) a system of notation could only be defensible, it

[3] Such a result would not have to be the death warrant of programmes—the most famous is H. P. Grice's—in which the notion of individual speaker's meaning is assigned priority. But it would require such programmes so to construe individual speakers' meanings as to build in third-person availability from the outset.

seems, if it could be shown that *any* sort of rule-following was impossible for such an individual. And that seems preposterous: there is no limit to the kinds of behaviour in which such an individual might engage which would be utterly inexplicable unless construed as involving purpose, insight, and the mastery of rules. (A nice example of Michael Dummett's: suppose Robinson finds a Rubik's Cube washed up on the beach, and learns to solve it. . . .)

A second constraint is similar but more immediate: 'going public', as it were, must make all the difference. There are two ways in which this constraint might be violated. First, a seemingly genuine problem might be disclosed in private language which, however, would remain even after sufficient alteration in the content of its vocabulary to enable a number of speakers reasonably to believe that they had the language in common. Obviously, whatever independent interest the problem might then possess, it could not provide the basis of a specific argument against private language. And the same holds, secondly, if, although going public would solve a specific difficulty disclosed with private language, some relevantly *analogous* difficulty could plausibly be argued to beset language in general. A cogent argument against private language must leave communal language alone. Abstractly so presented, the point might seem hardly worth stating. But it has been violated by a number of interpretations of Wittgenstein, favoured both by sympathizers and critics; and it should anyway alert us to the risk attending the reading of Wittgenstein in this context as any sort of epistemological *sceptic*, despite the temptations posed by passages like I. 265.

A third desideratum, if not quite a constraint, is that the argument should be effective against two quite different opponents. The usual antagonist, for most commentators, has been *Cartesianism* about sensations: a standpoint according to which not merely are sensations private to a subject, but they constitute necessarily self-intimating material—no aspect of whose character can pass unnoticed—about which the subject's opinions have an error-proof authority. Now, no doubt the privacy of sensation has some connection, in Cartesian thinking, with its putatively self-intimating and indefeasibly certain character. But that there is at any rate no entailment is evident in the possibility of a weaker standpoint—that of the *Fallibilist*—according to which the subject may be sincerely mistaken, or fall prey to oversight, in the description of his sensations, or other psychological material, its privacy notwithstanding. The

Cartesian conception encourages us to think of our sensations as events played out on a stage to which necessarily we alone are witness, but there is no reason why such a picture has to involve the idea that our witness has to be all-seeing and error-proof. Yet the Fallibilist standpoint—well represented in recent literature[4]—is as much committed to the possibility of private language: if you are privy, albeit fallibly, to events and processes of a kind of which you and I can never be *jointly* aware, even in principle, there is no means whereby I can make the sort of comparison between our respective linguistic behaviour necessary if I am to arrive at reason to think that we share an understanding of the language which we use to describe that material. Something worthy of the title of a demonstration of the impossibility of private language ought, therefore, to be effective against this weaker position. An argument of less general bearing, trading upon the additional features present in Cartesianism proper, might still be of some interest. But it would necessitate no more than repair to the Cartesian philosophy of mind; it would not extract the root. And it could not have the wider sort of philosophical significance outlined in section 2.

4. *Three Interpretations of I.258–60*

It is worth quoting in full G. E. M. Anscombe's translation of the passage with which we are especially concerned:

> 258. Let us imagine the following case. I want to keep a diary about the recurrence of a certain sensation. To this end I associate it with the sign '*S*' and write this sign in the calendar for every day on which I have the sensation.—I will remark first of all that a definition of the sign cannot be formulated.—But still I can give myself a kind of ostensive definition—how? Can I point to the sensation? Not in the ordinary sense. But I speak, or write the sign down, and at the same time I concentrate my attention on the sensation—and so, as it were, point to it inwardly.—But what is this ceremony for? For that is all it seems to be! A definition surely serves to establish the meaning of a sign.—Well, that is done precisely by the concentrating of my attention; for in this way I impress on myself the connection between the sign and the sensation.—But 'I impress it on myself' can only mean: This process brings it about that I remember the connection *right* in the future. But in the

[4] See, for instance, Ross Harrison, *On What There Must Be* (Clarendon Press, Oxford, 1974), Chapters 3 and 6; and Simon Blackburn, 'The Individual Strikes Back', *Synthese*, lviii (1984), 281–301.

present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'.

259. Are the rules of the private language *impressions* of rules?—The balance on which impressions are weighed is not the *impression* of a balance.

260. 'Well, I *believe* that this is the sensation *S* again'.—Perhaps you *believe* that you believe it! . . .

The principal question is: what is the thought at the end of I.258? Why should it follow, if 'whatever is going to seem right to me is right', that the very notion of the *correctness* of the private linguist's would-be descriptions is emptied of content? Wittgenstein's transition has, indeed, the appearance of simple question-begging against the Cartesian; for if infallible authority, about any subject matter, were ever a possibility, then—trivially—whatever seemed right to the authority would be right. The Cartesian conception is exactly that each of us is such an authority for the character of his own sensations. A reader could be forgiven the impression that Wittgenstein has done nothing to disclose a specific fault in that conception of *sensation*; his reservation, whatever it is, would apply, it seems, to any such putative cognitive authority. And, whether one intuitively suspects the coherence of the idea of such authority or not, it does seem unclear what exactly the reservation is.

It will be useful to consider a number of interpretations which contrast with the one I want eventually to recommend. There is, to begin with, the possibility of reading the passage along the following, broadly verificationist lines. Meaning is an essentially normative notion: if an action—like making a noise, or marking down a symbol—is to be credited with meaning, there has accordingly to be sense in the distinction between situations in which tokens of that action accord with this meaning and situations in which they do not. This is as much as to say that, from the point of view of a subject, trying to keep his actions in accordance with the meaning, there must be sense in the distinction between what seems right to him and what is right. For someone of verificationist sympathies, however, there has to be doubt about the *content* of a distinction which no one can possibly be in a position to exploit: a distinction, that is, between states of affairs which, in the nature of the case, cannot be verified to obtain independently of each other. For the verificationist, then,

there is, in the case of putatively private language, doubt about the sense of a distinction which *has* to make sense if the notion of meaning, and with it that of language, is to have any proper application.

The spirit of this train of thought could, of course, be formulated and refined in various ways. It is notable, in particular, that its purpose would be as well served by the slogan 'Meaning is Use' as by explicit versions of the verification principle. For it is exactly the lack of any contrastive *uses* for 'seems right to me' and 'is right' which Wittgenstein seems to be presenting as undermining any contrast in their content. I myself would not regard the presence of verificationist premisses as importing any error of substance. Indeed, I had better declare now that the interpretation I shall eventually recommend is not guaranteed to disturb the privacy of a theorist who is *sufficiently* resolute and comprehensive in his fidelity to the verification-transcendent. Nevertheless, the rather *direct* play which the kind of argument sketched makes with a quite general principle about meaning has to be associated with costs in point of pure cogency. The point is simply the unlikelihood that purely theoretical argument can provide absolutely compelling support for such a general principle: however well grounded it may seem, in the abstract, to be, consequences of a sufficiently *outré* character will generate suspicion. Anyone who wishes to base far-reaching metaphysical conclusions on (an analogue of) the verification principle had better recognize the substantial body of philosophical opinion which is prepared to find others of the known consequences of the principle to be sufficiently implausible to discredit its use as a tool of persuasion.

The answer provided by this first interpretation to our title question is, in effect: yes, provided there is a cogent argument for the verification principle, or for the principle that differences in meaning have to correspond to differences in use. What, I am suggesting, is doubtful is whether we have any conception of what it would be for argument in support of a principle of such generality to be *cogent* if 'intuitions' about the plausibility or implausibility of consequences are not allowed to come into play; and doubtful, in consequence, whether something as controversial as the present question can be resolved by an argument which relies directly on a general principle of this sort.

A second, quite natural interpretation finds Wittgenstein's thought to be that the private linguist has no basis for trust in his

own judgement, no effective controls on his own competence to practise the system of concepts which he believes his private language to encode. Now this is, I believe, a point that Wittgenstein is keen to have recognized; it may plausibly be read into I.265 and seems to be exactly the thrust of

> 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 (II, xi, p. 207).

It is, from the inner perspective of the private linguist, all the same whether his ongoing practice keeps accord with the original, putatively definitive use of '*S*', or whether his memory of the kind of sensation he wished to mark lets him down repeatedly and in diverse ways, so that his use of '*S*' is actually quite chaotic. And there is, of course, no *other* perspective to be had. But what exactly does the point show? That the Cartesian conception of the inner, characterized above, is itself unstable under sceptical attack. Wittgenstein is urging on the Cartesian the realization that there is a difference between a guaranteed *fit* between impressions and fact and the mere inability to distinguish them. Once it is insisted that there really is to be such a thing as a *correct* description, in the private language, of the subject's occurrent sensation, whose correctness is settled by the character of that sensation and the original private ostensive definition, there is absolutely no basis for an opinion about the reliability or otherwise of the subject's impression—so scepticism seems to be the only rational standpoint.

The point is well taken. But, whatever Wittgenstein's intention, it is plain that it is effective only against the Cartesian and does not tell against the possibility of private language as such. The Fallibilist counter will be to acknowledge that, ultimately, there is no decisive riposte to sceptical doubt about his private linguistic competence, but then to enquire how exactly 'going public' would get around the difficulty. For one thing, the subject will still need to rely upon his own memory, and other faculties, if he is to make effective use of others' judgements as a check on his own, since ultimately he must judge for himself what their judgements are and whether they accord or conflict with his.[5] For another, no sooner do we admit—as we

[5] A point first made, I believe, by A. J. Ayer in the symposium with Rush Rhees, 'Can there be a Private Language?', *Aristotelian Society Supp. Vol.* xxviii (1954), 251–85; see p. 256.

are inclined to do—that correctness is never *constituted* by a community-wide consensus than we leave scope for a similar kind of sceptical doubt about the shared judgements of the speakers of a public language: if coincidence between correctness and communal consensus is a matter of the proper functioning of individuals' faculties, it is difficult to see what could obstruct the possibility of sceptical query, of one or another familiar sort, about whether, in any particular case or range of cases, such proper functioning had taken place. If, on the other hand, there proved to be some sort of *conceptual* connection between correctness and the considered verdict of a whole linguistic community, that might well exempt the community from forms of sceptical doubt to which the private linguist is vulnerable; but it would be a conclusion for which we still await the supporting argument—which would have to embrace considerations quite outwith the scope of the interpretation we are presently considering. If this interpretation were correct, indeed, the verdict would have to be that, as a purported demonstration of the impossibility of private language, I.258–261 falls foul of *each* of the three constraints outlined above.

There is evidence, however, that Wittgenstein may have wished a would-be believer in the possibility of private language to ponder what is apparently a deeper-reaching sceptical doubt.[6] A little earlier (I.237) he writes

> Imagine someone using a line as a rule in the following way: He holds a pair of compasses, and carries one of its points along the line that is the 'rule', while the other one draws the line that follows the rule. And while he moves along the ruling line he alters the opening of the compasses, apparently with great precision, looking at the rule the whole time as if it determined what he did. And watching him we see no kind of regularity in this opening and shutting of the compasses. We cannot learn his way of following the line from it. Here perhaps one really would say: 'The original seems to *intimate* to him which way he is to go. But it is not a rule.'

The quotation reminds us that there is a difference between following a rule and acting under a sense of constraint; it is even possible, as a result of hypnosis or in the throes of who-knows-what sort of

[6] This line is briefly canvassed, if I read him correctly, in Leslie Stevenson's *The Metaphysics of Experience* (Clarendon Press, Oxford, 1982), 45–6.

manic episode, to have the impression that you are rule-following, that your performance is guided by a concept, when that is not the case. The practice of the private language is to be, at least, a rule-governed practice. What reason, then, can the private linguist offer *himself* for thinking that he genuinely has such a practice, that his performance is not to be compared to that of the lunatic with a pair of compasses?

This train of thought is in one way more, and in another less sceptical than that which we have just discarded. It is more sceptical because what is in doubt is not the *reliability* of the judgements of the private linguist but whether he is so much as making judgements at all; at issue is not what grounds there are to think that the private linguist is competent to follow certain rules but what grounds there are to think that there are any appropriate rules in the offing. But the train of thought is, at least arguably, also less sceptical inasmuch as it does not seem to need recourse to doubt about the subject's memory. Under the second interpretation, we were to imagine that changes in the type of sensation which the subject was willing to describe as '*S*' are compensated for by changes in his memory of the kind of sensation which he originally so baptized; but the delusion, whatever exactly its nature, suffered by the lunatic with the compasses need not involve, it seems, his misremembering anything. This asymmetry holds out the promise that the interpretation may serve better than its predecessor. Challenged to produce a reason for thinking that he is using '*S*' correctly, the private linguist will naturally refer to his memories of his previous practice; the challenge, if it is to continue, must then assume the form of a demand for reason to rely on those memories. But, challenged to give himself reason to think that he is so much as following rules at all, it is not clear that the private linguist has even the beginnings of a reply. If the lunatic need be under no *further* illusion about his previous conduct and experience, a debate (with his doctor, say) about whether he really is rule-following will have to involve considerations of a quite different kind. And now what could such debate pivot around except the capacity of the lunatic to explain, or of others to interpret, either the specific rule which he believes himself to be following or, at least, the *kind* of rule it might be—perhaps one too complex for others to follow, for example—and what was at stake for him in his pursuit of it? The emergent suggestion is, accordingly, that nothing counts as giving yourself, or anyone else, reason to take certain episodes of

your behaviour as genuine cases of rule-following which does not consist in supplying reason for thinking that it could reasonably be so interpreted *by others*.

This seems a promising outline, whose detail it may well be possible to fill in convincingly. If so, the private linguist does confront a distinctive difficulty: he cannot give himself even the weakest reason for thinking that he has a rule-governed practice. Robinson Crusoe, in contrast, need confront no such difficulty; at least, it is not an *immediate* consequence of his social isolation that nothing in the course of his experience, as he interacts with his desert island environment, can give him reason to think that others would reasonably regard him as rule-following if appropriately placed to attempt to interpret his behaviour. Naturally, he may have no defence against a *sceptical* doubt on the point; but the challenge to the private linguist, under this third interpretation of the argument, is not to produce sceptic-proof reason to think that he is rule-following, but merely to show that he is in no way disadvantaged by comparison with Robinson or the individual in the community. Since it seems clear, in addition, that it is merely privacy, rather than the assumption of any sort of Cartesian certitude, which gives rise to the difficulty, we appear to have the germ of an argument which, if successfully developed, might succeed in observing all the outlined constraints.

However, second thoughts are less encouraging. Even if we anticipate a completely convincing demonstration that the range of reasons to believe a subject to be rule-following is restricted in the way suggested, two major sources of dissatisfaction remain. First, accepting that no one can have even the weakest reason for supposing that the private linguist has a rule-governed practice, transition to the conclusion that it cannot be *true* that he does would seem to require mediation by some general form of verificationism, so that the cogency of the argument confronts the same impediment as afflicted the first interpretation. Second, whatever the extent of one's sympathies with verificationism, the fact is that there is a large class of propositions whose acceptance is normal among us, and whose non-acceptance, in very many cases, would be unthinkable, for which we nevertheless can supply no solid reason. Examples like 'There exist other consciousnesses besides my own', 'The world did not come into being five minutes ago', 'There are material bodies' all have the feature that they are *beyond* evidential support: that unless

they are presupposed, nothing counts as support for the sort of more specific claims from which they may be inferred. Simon Blackburn[7] makes a case for supposing that a similar groundlessness ultimately belongs to the belief that we share communal language; if that is so, the third interpretation does not, after all, fully satisfy the second of our earlier constraints. But whether Blackburn is right in detail matters less than the reflection that we do not in general wish to equate groundlessness with unacceptability (or, if we do, it is only human irrationality which explains why scepticism is not a dominant ideology). Accordingly, the groundlessness of the private linguist's belief that he has a rule-governed practice would not be *eo ipso* a criticism of that belief; it would remain to be shown that it is not a belief to which he is perfectly entitled, a framework belief, as it were, comparable in its own sphere to our belief in the existence of other minds or of material bodies.

Each of the three interpretations, even if not providing something fully cogent, nevertheless affords a useful and instructive version of the argument. Collectively, they suggest that a defender of private language had better be anti-verificationist, anti-Cartesian, and a proponent of the propriety of groundless belief. It immediately occurs to one to wonder whether this is a consistent set of characteristics: whether, in particular, an anti-verificationist can make a decent fist of explaining our entitlement to beliefs which, on his construal, are associated with transcendent truth-conditions for whose obtaining we can, in the nature of the case, have absolutely no evidence. One possible way of completing the demolition of private language would be to show that there is indeed an irresoluble tension here. Another would be to show that the province of legitimate but groundless belief cannot extend to the case in point. And a third would be, of course, further research into the semantic presuppositions of verificationism. In each of these various ways, then, the issue is very open. However I do not think that we have yet considered the most persuasive version of Wittgenstein's train of thought, a version which, although textual evidence can be found in support of each of the interpretations so far considered, arguably best fits the actual letter of I.258–61. Before we do so, let me briefly explain why I believe the interpretation which takes centre stage in Saul Kripke's *Wittgenstein on Rules and Private Language*[8] ultimately affords no cogent argument against the possibility of private language.

[7] Loc. cit., pp. 291 and following.

5. *Kripke's Interpretation*

A wide class of long-standing philosophical disputes pivots on the question whether the statements of some problematic region of discourse—ethics, aesthetics, theoretical science, and pure mathematics are the most immediate examples—should be regarded as having a *genuinely factual* subject matter, or whether—as argued for example by certain forms of ethical prescriptivism, and by scientific instrumentalism—their grammatical form contrives to mask the fact that their role is not that of fact-stating at all.

There are a variety of ways in which an anti-factualist standpoint might be supported. One way—essentially Hume's strategy with statements about causation—would be to argue that a preferred epistemology recognizes no means for our attaining any adequate conception of the putative species of fact. Thus there is, for Hume, within the means provided by his preferred empiricist epistemology, no way of attaining any satisfactory idea of the *necessitation* which the causal relation is supposed to involve; the way is then open—or so Hume believes—to regard that component as marking, rather, the projection of an attitude which human beings naturally take up towards tried and tested regularities. A second approach would be to attempt to establish a topic-neutral account of ways in which the distinction between fact-stating and non-fact-stating declarative sentences emerges in their uses, and then apply that account to the problematic class in question. The difficulty with this way of proceeding, of course, is that the general topic-neutral account has not merely to isolate marks of a distinction which we already perfectly understand but must, in addition, go some way towards legitimating the distinction.

There is, however, a rather ingenious third strategy. The leading idea is that any genuine species of fact ought to be knowable by an appropriately endowed subject. Accordingly, if we can specify what powers a subject would have to have in order to be in a position to know a fact of some putative species, and if it then emerges that even

[8] Blackwell, Oxford, 1982.

such a subject could make no such defensible knowledge claim (could provide no compelling reason for preferring one such claim to another incompatible with it), it must follow that there cannot be any such facts in the first place. It is this third strategy which informs the Sceptical Argument which Kripke finds at the heart of Wittgenstein's discussion of following a rule in I.185–242. The conclusion of the argument, as Kripke interprets it, is that all talk of meaning and understanding ceases to qualify as factual, so that there are no objective truths about, for example, what particular expressions mean, or how subjects understand them, or what specific uses accord with their meanings.

An argument against private language now emerges under the aegis of a non-factualist reconstrual of this region of discourse. It is this reconstrual which Kripke, following the Humean model, calls a Sceptical Solution. Sentences like 'Jones means addition by "plus"' and 'If Jones means addition by "plus", he will answer "259" when asked "What is 132+127?"' may still, according to the Sceptical Solution, possess a determinate use even if divested of objective truth conditions. And the respective uses which they have are, roughly, to ratify Jones's membership of the class of speakers whose uses of 'plus' are generally reliable, and to express a test for membership in that class which the speaker believes would be ratified by the responses of those already accredited with membership. The suggestion, in general, is that once we prescind from the idea that talk of meaning and understanding is apt to convey substantial matters of fact and look instead to an account of the role and purposes which such talk plays in our lives, we find that there is invariably a reference, explicit or implicit, to the linguistic practices of a speech *community*. Kripke's interpretation of the private language argument would then draw the conclusion that such a reference is essential to the legitimate use of those concepts, and that they can accordingly have no proper application to the activities of a would-be private linguist.

The foregoing is just the barest outline of Kripke's interpretation: it can convey little of the excitement and fertility of Kripke's book to someone unfamiliar with it. But it will serve to indicate various points at which, in my judgement, Kripke's version of the private language argument fails of cogency. To begin with, the whole package can succeed—obviously—only if the Sceptical Argument succeeds, and it is open to serious question whether it does so; I have,

however, already had occasion to try to indicate places where the Sceptical Argument seems to me to go astray, and I have no space to elaborate on that discussion here.[9] There are, however, two slightly less obvious points. First, even if the success of the Sceptical Argument were assumed, much more would need to be done to explain why the Sceptical Solution is not just an optional extra. Hume *could* quite consistently have drawn the conclusion that we should simply drop the notion of causation as a piece of discredited mythology. Kripke writes as though a parallel option was not available in the case of meaning, since it would involve us in the 'incredible and self-defeating conclusion that all language is meaningless'.[10] But that cannot be the right way of formulating the conclusion of the Sceptical Argument, since it involves, in effect, descriptive use of the very notion whose grip on reality the Sceptical Argument aims to dislodge. It is rather as if we construed the conclusion of Hume's argument as being that all regular associations between events are *coincidental*, when the point is rather that nothing in the world—if the argument is sound—corresponds to our distinction between causal and coincidental regularity. The conclusion of the Sceptical Argument must—at least if it is to be a factual claim—be something expressible without recourse to any notion cognate to that of meaning; so it cannot be the 'incredible and self-defeating' conclusion described by Kripke.

Quite how it should be formulated is a question which we can leave to be pondered by proponents of the argument. For present purposes it suffices merely to note that it is at least not obvious that the conclusion has to take a self-defeating form; and that what would be called for would presumably be something akin to the Quinean programme for an account of language, and language-related institutions, in which the traditional notion of meaning plays no part. It is accordingly an *assumption* that the Sceptical Solution is so much as called for, and any conclusion about private language, drawn in the manner indicated, can so far at best be provisional. The argument needs a back-up demonstration either that no Quinean

[9] See 'Kripke's Account of the Argument against Private Language', *Journal of Philosophy*, lxxxi (1984), 759–78. For additional criticism of the way Kripke involves the community in the Sceptical Solution, see Warren Goldfarb's critical study of Kripke's book, in the same journal, lxxxii (1985), 471–88.

[10] Kripke, loc. cit., p. 71.

account of language can be satisfactory or that such an account would independently be inhospitable to private language.

It is in any case open to question whether the Sceptical Solution is so much as coherent. The problem is that the notion of meaning is platitudinously connected with that of truth: whether an utterance expresses a truth is a function only of its content and the state of the world in relevant respects; and the content of the utterance is, in turn, a function of the meanings of its constituents, the way in which they are therein put together, and the context. So if there were no facts about meaning, it would appear to follow—by the compelling principle that non-factuality among the parameters in a question must divest the question of factuality too—that there can be no facts about an utterance's truth value either. Such an outcome would call into question the very possibility of explaining what exactly it is that the non-factualist about a given region of discourse intends to hold.

There is no doubt a great deal of scope for further exploration of the implications of and possibilities for Kripke's interpretation. But perhaps enough has been said to justify the claim that, whatever else such exploration might teach us, it is unlikely to disclose the cogent argument which we seek.

6. *The Recommended Interpretation*

The verificationist interpretation appealed, in effect, to the idea that the independence of the requirements of a rule from what anyone takes them to be is a precondition of genuine rule-following; so that the need for a distinction between what seems right to the private linguist and what really is right is a simple consequence of the consideration that private language, like all language, is to be a rule-governed practice. But the trouble is that exactly this platitudinous-seeming conception of the autonomy of rules is one of the things which is under scrutiny in the discussion of rule-following in the *Investigations* and *Remarks on the Foundations of Mathematics*. A typical passage (RFM IV.48):

> Then might it not be said that the *rules* lead this way, even if no-one went it? For that is what one would like to say—and here we see the mathematical machine, which, driven by the rules themselves, obeys only mathematical laws and not physical ones. I want to say: that the working of the mathematical machine is only the *picture* of the working of a machine. The rule does not do work, for whatever happens according to the rule is an interpretation of the rule.

I do not believe that Wittgenstein intended to question the *correctness* of the platitudinous conception: his intention was rather to expose certain misunderstandings of it to which he believed we are prone. But we would need to be sure, it seems, exactly how matters stand with the autonomy of rules before the earlier verificationist interpretation of the private language argument could be considered cogent, even if the correctness of some appropriate form of verificationism were assumed. Fortunately, we can bypass the issue. There is a different, somewhat more subtle route to the perception that a 'seems right'/'is right' distinction is prerequisite for the private linguist's enterprise, which there is some textual evidence Wittgenstein may have had in mind.

The suggestive passage is the exchange in I.260 quoted above:

> 'Well, I *believe* that this is the sensation *S* again.'—Perhaps you *believe* that you believe it!. . .

Wittgenstein's response to his interlocutor has no point unless he thinks that he—the interlocutor—may believe that he has such a belief yet be wrong. But in what circumstances can a second-order belief—a belief about one's beliefs—be incorrect? The sort of cases that most immediately spring to mind would involve the conception that the true character of one's (first-order) beliefs can be veiled in the sort of way that notions like self-deception, or the psychoanalytic idea of the unconscious, standardly presuppose. But that sort of example is not to our purpose. A quite different possibility—the one which, I suggest, Wittgenstein is getting at—is that *X*'s belief that he believes that *P* is false because there is *no such thing as* the belief that *P*. That will be the situation whenever the clause apparently specifying the propositional content of the supposed first-order belief, although prima facie apt for that role, actually fails to have the appropriate kind of content. A genuine belief is an attitude to the type of state of affairs described by the clause which specifies the content of the belief, namely the conviction that it is realized. So if there simply is no such thing as that type of state of affairs (which is not, of course, the same issue as whether there exists a particular state of affairs of that type), no more can there be any such belief. Accordingly, if the clause purportedly supplying the content of the belief belongs to a family of declarative sentences which correspond to no genuine matters of fact, it cannot be used to specify the content of a possible belief. (You can no more believe something which is

not apt to be genuinely true or false than you can wonder what is the answer to something which is not really a question.)

On this interpretation, then, Wittgenstein's suggestion is that the existence of a genuine distinction between what seems right to the private linguist and what is right is necessary if the private language is to be apt for the making of *genuine statements*. And so there would be a point of affinity with Kripke's discussion. Only Wittgenstein's point would not be that talk of meaning and understanding *in general* lacks a factual subject matter, but that the 'sentences' of a private language cannot qualify as having a factual subject matter unless the 'seems right'/'is right' distinction can be made good for them—which, he has suggested in I.258, it cannot.[11]

Argument is now needed for two claims: first, that there is indeed the connection between the 'seems right'/'is right' distinction and factuality which the interpretation demands; and second, that the distinction eludes the resources of the private linguist.

The grounds for the first claim are relatively straightforward. What, more exactly, is at issue is whether the sentences of a given family are apt for the expression of fact only if:

(a) *X* believes what '*P*' expresses

and

(b) What '*P*' expresses is true

have an appropriately contrasting content where '*P*' is any of (appropriately many of)[12] the sentences in question. Now, it is in the nature of facts to stand to us in various cognitive relations—like

[11] What will follow, if Wittgenstein is correct, is not, strictly, that private language is impossible, but that it cannot provide a medium for the formulation of genuine statements, commands, questions, wishes, the framing of hypotheses, or any kind of speech act which presupposes the availability in the language of the means for depicting genuine states of affairs. It is a further question whether anything so impoverished as to lack all these expresive resources could qualify as a language (perhaps a private language apt only for the expression of expletives would still be a possibility). However since all the lines of thought which attract, or pressure, towards the possibility of private language involve regarding it as a medium for expression of knowledge, there is no comfort for anyone in such a possibility—if possibility it be.

[12] There are very delicate questions here concerning what, if any, notion of a *family* of sentences would sustain the validity of the principle that if some members of a family are factual, all are; and whether every member of such a family should be required to admit the (a)/(b) contrast; and who *X* may be. Presentation of the argument which follows does not require us to engage them—though a supporter of the Blackburn/Harrison strategy (see below) would have a definite interest in doing so.

being known, overlooked, reasonably supposed, and so on. A realist (contrast: verificationist) may press us to countenance in addition a class of facts which are cognitively inaccessible, about which neither knowledge nor reasonable opinion can be achieved. Whatever the merit of that thought, it at least suggests an argument by dilemma. If '*P*' is apt for the expression of fact, then there must at least be sense in the idea of a subject standing in one of the various possible cognitive relations to facts which are relevant to the appraisal of '*P*'. So if—going along with the realist—all such facts are inaccessible, *X* can only be ignorant of them, and any belief he has about the status of '*P*' can have no bearing on the likelihood of its truth. If, on the other hand, certain relevant such facts are accessible, the cognitive processes involved in their investigation must nevertheless be *fallible*, and hence any opinions formed by *X* as a result of such investigation can be at best a fallible guide to the likelihood of the truth of '*P*'. Either way, there will be no entailment from (a) to (b); either way, then, there has to be an appropriate contrast in their content.

An immediate afterthought is that this reasoning begs the question against the Cartesian. The Cartesian conception is exactly that a subject's sensations, for instance, constitute a domain which is transparent to him, about which his judgements cannot be mistaken; so if '*P*' and the other sentences in the family are apt for the description of such a domain, the entailment from (a) to (b) must go through. Admittedly, we have already recorded a doubt, in discussion of the second interpretation above of Wittgenstein's argument, whether the Cartesian account can defend itself against scepticism. But that would not exonerate the present interpretation from the charge of question-begging. It would be better to show that the Cartesian is committed to upholding a contrast in content between (a) and (b) for a different reason.

Well, that can be shown quite easily. For the Cartesian claim is precisely that *X* possesses certain *guarantees* when it comes to judgement concerning the character of, for example, his own sensations which he does not possess elsewhere. This claim has to be understood as comparable to the claim, for example, that the arithmetical functions, x^2 and $8(x-1)-4$, although divergent in value generally, necessarily coincide in value when $x=6$. There is no way of understanding what the Cartesian believes is guaranteed to *X* unless we take it that the sentence forming operators, '*X* believes what——expresses' and 'What——expresses is true' have their standard

senses. Any substitution, within their gaps, of a sentence with a factual subject matter about which *X* has no Cartesian authority will reveal, by the argument just run through, that those senses are different. Hence the Cartesian has to grant that there will be some sort of difference of sense in the sentences which result respectively from substitution of a sentence about whose subject matter *X does* have Cartesian authority, notwithstanding the mutual entailment which he then wishes to claim.

So much, then, for the first claim. The second claim is that such an 'appropriately contrasting content' cannot be given to (a) and (b) if '*P*' is a sentence of a putative private language. But what exactly should count as the existence of the appropriate kind of contrast? The Cartesian at least is in no position to advert to a difference in the *truth conditions* of sentences corresponding as (a) and (b).[13] And a moment's reflection discloses that the obvious thought—namely that (a) and (b) differ in content merely because the sentence-forming operators which they respectively involve are different—is merely question-begging. For that consideration suffices for them to differ in content only if '*P*' expresses an argument appropriate for those operators to be applied to, that is, expresses a thought apt to be believed or true—the very point at issue.[14] So how should matters proceed?

The least that is required, it seems, is that the *information* conveyed by (a) should differ from that conveyed by (b). And it is plausible that, for a large class of examples, two items of information differ just in case there can be such a thing as reasonably regarding oneself as possessing one without the other. It is in exactly this sense that the informational content of two logically equivalent sentences may diverge, since unawareness of a logical equivalence need not be unreasonable. So let us propose as a general, though perhaps not exceptionless, principle that items of information coincide for a subject just in case to have reason to accept either, in a sense which need not involve closure under logical implication, involves, for him, having reason to accept both. Now there ought, presumably, to be no objection to supposing, of any particular belief held for good reason, that the subject is aware that he has reason for that particular belief. (That is: if there is no conceptual difficulty in the idea of a particular subject having reason for a particular belief, it cannot *introduce* a difficulty to suppose that he is aware of the fact.) It follows that—unless they are to be exceptions for the principle—we are entitled to regard (a) and (b) as conveying different items of information only if someone could have reason to believe one but not the other and could be aware of the fact. The argument will now be that when '*P*' is, putatively, a sentence of a language which no two people can reasonably believe they share, that is not a possibility.

[13] How far the Fallibilist, on the other hand, might get with a purely truth-conditional construal of the (a)/(b) contrast is a question for section 10.

[14] I do not mean to deny that '*X* believes what "*P*" expresses' and 'What "*P*" expresses is true' have different meanings even if '*P*' is not so apt, or is not a declarative sentence at all. But the Cartesian requires, in addition, that they be capable of simultaneous *truth*. It is that which the 'argument' from difference of composition cannot, unsupplemented, give him.

It will suffice to consider four cases. Letting '*A*' range over believing subjects, we have

(i) *A* is aware of possessing both reason to believe (a) and reason to doubt (b);
(ii) *A* is aware of possessing both reason to doubt (a) and reason to believe (b);
(iii) *A* is aware of possessing both reason to believe (a) and no reason to believe (b);
(iv) *A* is aware of possessing both reason to believe (b) and no reason to believe (a).

Our question is whether any of (i)–(iv) represents a possible conscious state of information either for an *A* who is some *Y* other than *X*, or for *X* himself. But no *Y* distinct from *X* can, it appears, enjoy any of the four states. Consider states (ii) and (iv). Both involve *Y*'s possession of reason to believe that '*P*' expresses a truth. But he cannot have reason to think that he knows *what* truth it expresses since, by hypothesis, '*P*' is a sentence of a language private to *X*, a language which no one could have reason to think he shares with *X*. So the only ground *Y* could have for supposing '*P*' true would be *X*'s authority on the matter—which contradicts the supposition that *Y* has reason to doubt (a), or at least no reason to believe it, which (ii) and (iv) respectively involve.

State (i) involves *Y*'s possession of reason to doubt that '*P*' expresses a truth despite accepting that *X* believes that it does: that would require either that *Y* has some independent purchase on what

'*P*' says—contrary to hypothesis—or, at least, that he has reason to think that some third party does who does not share *X*'s view and in whom *Y* places greater trust. But what reason could *Y* have for thinking that *X* and the third party share an understanding of '*P*' when the hypothesis of privacy dictates that the third party could have no such reason?

Y's enjoyment of state (iii) might seem less problematic. If, after all, *Y* has no conception of the meaning of '*P*', he can have no reason to think either that *X* is, or that he is not, more likely than not to be reliable about the sort of subject matter with which it deals. Hence he has no reason to regard grounds for thinking that *X* believes '*P*' to express a truth as tending to support (b). But one trouble with this thought is that it illicitly equates the supposition that the sense of '*P*' is private to *X*—that is, that it is impossible for *Y* to have reason to think that he shares an understanding of '*P*' with *X*—with its being impossible for *Y* to have *any* conception of the kind of subject matter that '*P*' deals with. That is not an equation which can be endorsed by the defender of the possibility of private language about the psychological. It is, for example, supposed to be consistent with my recognition that I can know neither the specific character of your sensations nor (in full) the meaning of the language in which you describe them that I can nevertheless apply the concept of sensation to you and hence have a general idea of the kind of thing which, when you give voice to the character of your sensations, you seek to describe. Generally, if one may at least suppose that '*P*', as used by *X*, concerns some aspect or other of *X*'s psychological state, then one is bound to take reason for (a) as supporting (b); and note that the point is not dependent on crediting *X* with Cartesian authority for his psychological states—it is enough that he be any sort of (fallible) authority about them, that his opinions about them count for something.

In section 2 I noted a different kind of pressure for admitting private language, exerted by sceptical predations on the communicability of intention in general and hence of idiolectic semantic rules in particular. If the possibility, indeed the universality of private language were granted for that sort of reason, there would not, presumably, be the sort of difficulty with (iii) just described. But there is a different, prospectively insuperable difficulty, connected with an aspect of the holistically interlocking character of psychological concepts. Without attempting detailed argument, it is at least highly

plausible that it is only on the hypothesis of a subject's possession of certain intentions that his linguistic or non-linguistic behaviour can be regarded as expressive of a certain belief. The relevant sort of example is familiar: dashing out into a street full of traffic may evince the belief that there is space to cross safely to the other side, if we assume, say, that the subject intends to make an appointment in fifteen minutes' time; but the same action may also evince the belief that there is no such space, if we assume the subject's intention to be suicide. The point, then, is that if we really took seriously the idea that all intention is inscrutable from a third person's point of view, we should surely be committed to saying the same about belief; and case (iii) would be precluded not because reason to believe (a) would give *Y* reason to believe (b) but because he could not have reason to believe (a) in any case.

Let us try matters from *X*'s own point of view. We can immediately discount (i) and (ii). If *X* were to enjoy the state of information described by (ii) he would be in a position to claim '"*P*" is true but I do not believe it'—precisely Moore's 'Paradox'—and state of information (i) would put him in a position to claim the equally paradoxical obverse: 'It is not the case that "*P*" is true but I do believe it'. Whatever the correct analysis of utterances of this bizarre sort, it is clear that they are not apt for the expression of self-consciously held reasonable belief and so do not correspond to possible states of information in the sense which concerns us. Further, it is not clear that the distinction between (iv) and (ii) can now be upheld: at any rate, if it can, it is because it is possible for *X* to be aware that he has no adequate reason for supposing that he has a particular belief without thereby acquiring reason to think that he lacks it. No doubt there is scope for further discussion of the matter, and it would be unwise to be dogmatic that the separation cannot be affected in certain sorts of (psychoanalytical) case. But it seems a dangerously nice distinction on which to rest a defence of private language.

A further point about (iv) emerges when we consider (iii). What the latter seems to describe is a state of information which *X* would occupy if he found himself smitten, as it were, with a belief for whose truth he recognized he had no adequate ground. Isn't that a possibility? Well—again—not if the subject matter of the belief is such as to render *X*'s very possession of it a ground for supposing it to be true; so not if the private language is to have a psychological subject matter about which a Cartesian, or weaker authority is claimed. But

a more general consideration is that, independently of the subject matter of '*P*', (ii) does not describe a state of information which a *rational X* can achieve for arbitrary '*P*'. No doubt restriction would be necessary on the thesis that a rational subject believes only what he has reason to believe. But this had better be true in general. The result is that a proponent of private language ought to regard *X*'s idiolect as containing the resources for expression of a wide class of statements for which (iii) can represent, for *X*, a possible state of information only if *X* is imperfectly rational. The same point applies to (iv). But the possibility of (iii) or (iv) for an imperfectly rational *X* had better not be the *only* thing that a proponent of private language can oppose to the claim that the needed contrast cannot be made between the content of (a) and (b); otherwise the conclusion will stand that a private language could not be employed by a perfectly rational subject—an interesting enough conclusion, and quite sufficient seriously to damage any philosophical thesis about language, or mind, which demanded the possibility of private language.

7. *Does This Interpretation Meet the Constraints on Cogency?*

So much by way of outline is the interpretation of I.258–60 which I wish to recommend. I am not claiming that Wittgenstein had exactly these thoughts but then mysteriously chose to put no more than a do-it-yourself kit on paper. What I claim for this interpretation is that it elaborates the *kind* of doubt which Wittgenstein's text seems to be urging us to have about private language, and that it promises a cogency missing from the other lines of interpretation considered. One thing necessary, in order to make good the latter claim, is to consider how the interpretation fares in relation to the constraints listed earlier, which is the task of this section. Then, in the next two sections, I shall review two lines of objection which force interesting further extensions of the argument.

First, the Crusoe constraint. Let '*P*' be any statement, of whatever subject matter, in Crusoe's invented language. It would appear, taking Crusoe as *X*, that he is as badly placed to draw a distinction between (a) and (b) as the private linguist, and for exactly the same reasons: none of (i)–(iv) represents a state of information which he can achieve if he is entirely rational. It is also plausible that, so long as Crusoe remains isolated on his desert island and no one else has the slightest idea of what, if anything, any of his utterances means, no third party *Y* can achieve any of the states of information (i)–(iv)

for the reasons described above. (I say 'plausible' because at least one of the considerations relevant to the discussion of private language, namely sceptical pressure towards third-person inscrutability of intention, would not be relevant.) Accordingly, let us suppose that Crusoe's isolation *de facto* prevents anyone else from arriving at any of the states of information whose possibility is demanded by a distinction in content between (a) and (b). We can even suppose Crusoe's isolation to be permanent and unalterable. Still, the difficulty about concluding that the argument, if sustained, defeats Crusoe's language too is that we do not ordinarily suppose that it suffices for a distinction to be unreal that nobody is *in fact* going to be able to make any use of it. The crucial matter is, putting it crudely, whether we know *what it would be, as rational subjects, to have occasion to make use of the distinction*. No doubt clarification is desirable of what exactly such knowledge should be held to consist in. But, whatever the truth about that, it is surely highly questionable a priori whether there can be such a thing as knowing what, counterfactually, it *would be* to be in a certain state of information if that state of information *necessarily* cannot be achieved. The thought governing the whole argument was that (a) and (b) possess an appropriately contrasting content only if the information which they convey is different; which it is only if there *could be* such a thing as having one item of information without the other. Any reader who found the principle plausible will not have supposed that the 'could' was meant to be constrained by contingent obstacles to investigation; but if it is not constrained by conceptual obstacles, it is doubtful whether it can mean anything at all.

There were two aspects to the second constraint, that 'going public' must remove the difficulty. First, the precise difficulty besetting the private linguist must disappear; and second, there must be no analogous difficulty for public language. Now, the precise difficulty is that, when *X* is taken to be the private linguist, neither he nor anyone else can attach appropriately contrasting informational content to (a) and (b). But 'going public' means: bringing it about that '*P*' has sufficient public content to enable others, at least in favourable circumstances, to form a well-founded opinion about its truth value independent of *X*'s opinion. That ought to make space for (iv). Of course *X* may still be invested with authority about the truth value of '*P*'. But the possibility of independent assessment is likely to involve that that authority is, if only in special circumstances,

defeasible—which is to say that there ought to be space for some or all of (i)–(iii). In any case, whether or not the feasibility of (i)–(iv) follows from the very publicity of the meaning of '*P*', the important point is that there is never any significant obstacle to establishing the (a)/(b) distinction for the statements of a public language. What counts as enjoying one of the states of information, (i)–(iv), varies depending on the subject matter of '*P*'; but there is in general no difficulty in explaining what it would be to enjoy one of those states of information *vis à vis* some third party *X*.

Whether the second aspect of the 'going public' constraint—the demand that there be no analogous difficulty for a public language—is satisfied may seem less straightforward. Suppose we take *X* as the entire linguistic community. Then may not the argument for the case $A = X$ proceed more or less as before? Analogues of Moore's Paradox and its obverse may as well be formulated using 'we' rather than 'I'; and it is presumably legitimate to hypothesize that the community is collectively rational, so that at least the final consideration urged against the possibility of (iii) and (iv) in the relevant part of the earlier discussion can survive. As for the situation from the point of view of *Y* distinct from *X*, the fact is that there is, by hypothesis, no such *Y* who understands the language. So the argument can proceed much as before with, perhaps, only the additional consideration, for the case of (iii), that confrontation with a solid communal consensus about the truth of a sentence would have to provide an observer with *some* reason to suppose that sentence true even if he had not the slightest idea what it meant.

The thought is, in effect, that the argument outlined has no special bearing on private language but may be run for any entire linguistic community, whether of one or many members. It is a nice question what the proper response to the argument would be if that were so. But if we were right to conclude that the Crusoe constraint is satisfied, it cannot be so. And, in fact, the relevant consideration here is the same: the existence of contingent barriers to anyone's achieving one of (i)–(iv) provides no decisive basis for the conclusion that no one understands *what it would be* to achieve one of those states of information. The community will no doubt have a conception of the circumstances in which it would subsequently be forced to judge that such a consensus about the truth of '*P*' *had been* mistaken. So we know *now* what it would be for there to be a *Y* who was in a position to make that judgement about us: such a *Y* would be an additional

member of the linguistic community who had adequate reason for supposing that the circumstances had indeed induced mass error. There is, by hypothesis, no such *Y*, but there could be—just as there could be a translator of Crusoe's language. Accordingly, there can be no special difficulty with this aspect of the second constraint if it was right to conclude that the argument satisfies the Crusoe constraint.

The third constraint, conceived for the special case of the psychological private language on which Wittgenstein himself concentrates, was that the argument should be effective against both Cartesian and Fallibilist opposition. It will suffice to remind the reader that, by and large, the argument proceeds without special assumption concerning the subject matter of the private language or the authority, if any, about the truth value of his utterances supposedly possessed by the private linguist; and that at both points where such authority is presented as contributing towards the difficulties, it is of no consequence whether it is thought of as Cartesian or fallible.

Arguably, then, all three constraints are satisfied. Since he can make out no (a)/(b) contrast—either truth-conditionally; or by appeal to difference of composition; or by appeal to difference of informational content, in accordance with the principle of section 6—and since the reasons for demanding the contrast seem absolutely compelling, we discharge the Cartesian from further punishment at this point. But the Fallibilist is not yet done for.

8. *The Return of the Community*

One important line of objection to the argument is suggested by the thought, three paragraphs above, that even a community-wide consensus may legitimately be regarded as mistaken *at another time*. Moore's Paradox and its variant depend upon present-tense formulation: there is nothing bizarre about for example '"*P*" was true but I did not believe it'. But nothing in the original argument, that (a) and (b) must possess contrasting content if '*P*' is to count as apt for the expression of fact, seems to require the present-tense formulations which (a) and (b) actually assumed—the fallibility of our cognitive powers would be as well reflected by a contrast in the content of corresponding past-tense formulations. Such reformulation would not, to be sure, affect the course taken by the argument for the case $A = Y$, distinct from *X*. The objection, rather, will be that the discussion of the case $A = X$ was at best incomplete: we have simply

not considered, in particular, whether the private linguist could get in position to pass principled *retrospective* verdicts of ignorance or error on himself. Should it emerge that he can, the argument would be met head-on: it would be possible to grant its claims about what is necessary for factuality, while contending that the thought which Wittgenstein expressed by 'Whatever is going to seem right to me is right' in I.258 is just mistaken.

The question, then, is whether X, the private linguist, can himself assign an appropriately contrasting content to:

(a)* X believed what 'P' expressed

and

(b)* What 'P' expressed was true.

Let the private language be one for the recording of sensation types; and let P be 'This sensation is S'. No doubt it may happen that X is afflicted with a sensation which, at first, he wishes to record as an occurrence of sensation S but which, as it continues, comes to seem to him significantly different. He may then want to say with respect to some recent past time that (a)* is true but (b)* is not. Now that, so far, is not good enough. With what right does X propose to regard his former impression of the sensation as having been mistaken, rather than view the sensation itself as having subtly changed from an instance of type S to something different? If the latter were the correct account, both (a)* and (b)* should be allowed to stand. Hence there can be, for X, such a thing as a state of information warranting assertion of (a)* but denial of (b)*—a state of information of type (i) with respect to (a)* and (b)*—only when he has equipped himself with a criterion for marking off the irrelevant kind of case when it is the sensation itself which has undergone subtle change. But it is utterly unclear what form such a criterion might take: all X has to go on, it appears, is his own conscious phenomenology; and phenomenologically there simply *is not* any difference—things will seem the same 'on the inside' whichever of the two descriptions is supposed to be accurate.

The point is vivid in the case of sensation, but in fact it is of general bearing. More needs to be done by a proponent of private language, whatever the envisaged subject matter, than simply to invoke the possibility that the private linguist might find himself inclined to correct a former judgement. We considered two ostensibly contrasting accounts of the situation in respect of sensation S. We could in fact have considered three, the third being that it is X's inclination at the later time not to regard his then-occurring sensation as being of type S which is astray. So: on the basis of what principle should X give his present descriptive inclination priority over the previous one? This question must arise even if, for whatever reason, the subject matter of the private language is not conceived of as giving rise to the possibility of the kind of gradual transformation which would allow both of X's descriptions to stand. What needs to be shown by someone who would have us believe that a retrospective analogue of (i) is accessible to the private linguist is that a situation may arise in which it is *rational* to accept both (a)* and the negation of (b)*; that has not been accomplished if the contrasting uses envisaged for (a)* and (b)* are based on nothing more substantial than the subject's arbitrarily giving priority to his present over his previous descriptive inclinations.

There is, as far as I am aware, only one line of thought extant in the literature which promises a reply. This is Simon Blackburn:

> ... in the usual scenario, the correctness or incorrectness of the private linguist's classification is given no consequence at all. It has no use. He writes in his diary, and so far as we are told, forgets it. So when Wittgenstein imagines a use made of the report (e.g. to indicate the rise of the manometer) he immediately hypothesises a public use. He thereby skips the intermediate case where the classification is given a putative private use. It fits into a project—a practice or technique—of ordering the expectation of the occurrence of sensation, with an aim at prediction, explanation, systematisation, or simple maximising of desirable sensations. To someone engaged on this project, the attitude that 'whatever seems right is right' is ludicrous. System soon enforces recognition of fallibility.[15]

Blackburn does not elaborate, but it is not difficult to imagine the kind of thing he could have in mind. The essential idea is to let the private linguist become a *theorist*, concerned not just with the recording of sensations, say, but with anticipation of their future character and the disclosure of patterns of occurrence among them. And the thought is that, once he embarks on this theoretical project, the routine constraints on all scientific theorizing (of comprehensiveness,

[15] Blackburn, loc. cit., pp. 299–300. Compare Ross Harrison, loc. cit., p. 161. And see Ralph Walker's *Kant* (Routledge and Kegan Paul, London, 1978), 115, for a terse gesture at the same idea.

predictiveness, simplicity, and so on) will supply a framework in terms of which it may be rational, on occasion, to discount former, or even present, descriptive inclinations.

Let us construct a simple example. Suppose the private linguist is afflicted, as it seems to him, with three distinct sensation types which within any instance of some specified period—one hour, say—invariably comply with the following two hypotheses:

H_1: if no instance of S_1 occurs, then an instance of S_2 occurs:
H_2: if an instance of S_3 occurs, then no instance of S_2 occurs.

Imagine now an occasion on which the private linguist sets himself to record the occurrence of his sensations over an instance of the specified period. We can list the various possibilities in the form of a truth table:

	S_1	S_2	S_3
1*	T	T	T
2	T	T	F
3	T	F	T
4	T	F	F
5*	F	T	T
6	F	T	F
7*	F	F	T
8*	F	F	F

Line 1, for instance, is to be taken, in the obvious way, as expressing the possibility than an instance of each of the three types of sensation occurs during the specified period. The reader will note that the possibilities expressed at lines 1, 5, 7, and 8 are at variance with the two hypotheses, while the remaining four possibilities are consistent with them. Now, the general point is that although, ultimately, theory must be the slave of observation, a theory can fare sufficiently well over a sufficient period of time to make it rational to discount a prima facie discordant set of data. H_1 and H_2 may together constitute a repeatedly corroborated, long-standing little theory; or they may represent the output, for these three sensation types, of a well-entrenched theory of wider content. Either way, it can be rational for the private linguist, if at the end of the specified period he finds that his record is of one of the four types at variance with the theory, to look askance at his 'observations' rather than at the theory.

Let his policy in such circumstances be the commendable one of making the simplest adjustment to his findings consistent with retention of his faith in the theory and in his own powers to, by and large, record his sensations properly. And now ask, what, if that is his policy, ought he to say in each of the four possible cases of discordance? The answers are interesting. Clearly he will not make two adjustments if one will do. But if he has recorded occurrences of all three sensation types, that is, we are concerned with line 1, there are two possibilities open, namely rejecting that an instance of S_2 occurred, so transforming the data into line 3, and rejecting that an instance of S_3 occurred, so transforming the data into line 2. And *two* choices is one too many. The objection to which the introduction of theory was a response was precisely that the private linguist had no principled basis for according priority to a present preferred description over a former one. But, in the situation envisaged, a preference for his verdict about S_2 over that about S_3 or vice versa is open to just the same reproach. The lesson is that the objection is met only when there is some *uniquely* best way to resolve the inconsistency between theory and data; so long as there are choices, the decision among which is arbitrary, the invocation of theory has, in the present context, no point.

Matters are no better with line 8, when it seems to the private linguist as though no sensation of any of the three relevant kinds has occurred; either of the suppositions that S_2, or S_3, did in fact occur (unnoticed?) transforms the data into an acceptable set, and no non-arbitrary basis for a choice between them presents itself. But it is different with line 5: here re-assessment of the verdict about S_3 represents the best, indeed the only suitable transformation. So it would appear that this theory does indeed provide for the possibility of a state of information in which the private linguist may reasonably deem the statement that an instance of S_3 occurred to be mistaken: the state of information is that constituted by the grounds for the theory itself and his other recorded data. A similarly favourable case is presented by line 7 for the statement that no instance of S_1 occurred. The Blackburn/Harrison idea is accordingly a plausible one: equip the private linguist with well-entrenched theoretical beliefs and he may very well get into situations in which he may rationally discount verdicts which he would otherwise have endorsed and may indeed have endorsed for a time. Have we got private language off the ground?

It certainly is not as easy as that. Remember, to begin with, that the establishment of an appropriate distinction between (a) and (b) (or (a)* and (b)*) is only a necessary condition for qualification of a class of statements as factual. Suppose you found it possible, for example, to draw up an integrated system of moral principles in such a way as to coincide with the pattern of your untutored moral responses and to predict new ones. You might then have a theoretical basis, in just the sort of way envisaged, for describing certain of your moral judgements about particular situations as involving error or ignorance. But no non-factualist about morals need be dismayed by the prospect of such system; he will hold that it is one thing to talk as if moral judgements have a factual subject matter and another thing for them really to do so (though it would remain a decisive point in his favour if we could not so much as establish the practice of talking as if they did). We do not at present have an account of conditions both necessary *and* sufficient for a class of statements to qualify as genuinely factual; a proof of the possibility of private language would have to appeal to such an account.

Still, in fairness, Wittgenstein's claim was that a necessary condition could not be met. And Blackburn has at least suggested how it might be. But cause for dissatisfaction with the sort of theoretical scenario depicted emerges in any case, once we reflect on what further conditions are necessary for factuality. One compelling principle—(*T*)—is that a class of statements can qualify as factual only if every truth-functional compound of them does so; hence in particular the negation of any factual statement must itself be factual. But the theory described does nothing to establish the possibility of contrasting uses for (a)/(a)* and (b)/(b)* when *P* is taken as 'No sensation of type S_3 occurred'. So what the theory does achieve for 'A sensation of type S_3 occurred' is not enough.

In addition, the theory goes no way at all towards establishing the (a)/(a)*–(b)/(b)* contrast for either 'A sensation of type S_2 occurred' or its negation. A proponent of the Blackburn/Harrison strategy is, I suggest, going to be hard-pressed to explain why that sort of omission does not set up a kind of 'rotten apple' effect. There are two reasons for anticipating such an effect. The first, somewhat intricate, involves appeal to the principle—(*U*)—noted above in discussion of Kripke, that if the acceptability—to use a neutral term—of a statement is a function of that of others, some of which are deemed not to be genuinely factual, then the original statement may not be

regarded as genuinely factual either. The theory described, when well established, puts *X* in position to deny the appearance of acceptability of 'An instance of S_3 occurred' when and only when 'No instance of S_1 occurred' and 'An instance of S_2 occurred' are both acceptable with respect to the relevant period. Accordingly, if the latter does not qualify as factual, then—by (*U*)—neither, when asserted on that basis, does the conjunction of (a)* and the negation of (b)* with 'A sensation of type S_3 occurred' substituted for '*P*'. Hence—by the principle (*T*)—not both its conjuncts qualify as factual. Since it is completely unclear what ground sympathetic to a proponent of Blackburn could be given for regarding the first conjunct as failing in factuality, the trouble, it seems, would have to lie with the second: the statement, 'It is not the case that what "A sensation of type S_3 occurred" expressed was true'. And if that is not a factual statement, it ought to be permissible to say the same about 'What "It is not the case that a sensation of type S_3 occurred" expressed was true' and hence about the statement mentioned therein.[16] A further application of the principle (*T*) then gives the result that 'A sensation of type S_3 occurred' does not qualify as factual after all.

The second reason for anticipating a 'rotten apple' effect concerns whether the presence of statements in the situation of 'A sensation of type S_2 occurred' in the example described leaves open any possibility of a coherent account of the private linguist's methodology. No problem, of course, if all the ingredients in the private linguist's record of his observations may be regarded as factual: in that case the three observations represented by line 5, for instance, are straightforwardly *inconsistent* with the theory, retention of which will therefore demand some reappraisal of them. But it is not clear whether it is legitimate to invoke the notion of consistency in this kind of way if certain of the ingredients in a putatively inconsistent set are not genuine statements. I do not mean to dispute, of course, that perfectly respectable notions of consistency and inconsistency are applicable outwith the domain of genuine statements—to commands, for instance, and rules, as well as to moral judgements (if they are indeed

[16] The two principles implicit in this transition are:

(a) ⌜Not: what '*P*' expresses/ed is/was true⌝ is factual iff ⌜What 'Not: *P*' expresses/ed is/was true⌝ is factual;

and (b) ⌜What '*P*' expresses/ed is/was true⌝ is factual iff ⌜*P*⌝ is factual.

thought to be a doubtful case). The fact remains that the notion of consistency has its *primary* place when we are concerned with genuine statements. It is applied elsewhere, I suggest, only in a derivative sense. (Commands, for example, are inconsistent when and only when those corresponding statements are, which articulate the conditions under which the commands are complied with.) In any event, anyone who wants to apply the notions of consistency and inconsistency outwith the domain of genuine statements, and to sustain their ordinary connections with the notion of rationality, owes an account of what exactly inconsistency among the relevant sentences amounts to and why it is to be avoided. It is quite unobvious how such an account should proceed in the present case; certainly, nothing like the story appropriate for the case of commands can be appropriate here.

The foregoing remarks are premised on the assumption that unless the theory provides for the possibility, with respect to each particular 'observation' statement, of a state of information in which some sort of (a)/(a)*–(b)/(b)* contrast may be drawn, no sense has yet been attached to the contrast in that particular case. But someone may be inclined to question this assumption. In discussion of the Crusoe and publicity constraints above, it was contended that it suffices, in order for (a) and (b) to contrast in content, that we possess a clear account of what it *would be* to enjoy any of the states of information (i)–(iv); and that it is not required that any of those states of information *actually* be attainable. Now, cannot the private linguist claim to have such an account for each of the statements which he wishes to regard as a possible report of his sensations? For he would be in a position to make the appropriate contrast, in each such case, just when he happened to be impelled towards a theory which furnished a line in the truth table standing to that statement as line 5 stands to 'A sensation of type S_3 occurred' in our example. There is thus, it may be contended, no need to aspire to some *single* theory which can provide the needed contrast in every case; it is enough, for each putative report of sensation, to be hospitable towards the possibility of *some* theory's providing a basis for its correction.

This point, if sustained, would certainly be convenient for the private linguist. It would mean that he did not *actually* have to engage in any theorizing at all! It would be enough to indicate that he was 'hospitable' to the possibility of sometime attempting to integrate

his 'reports' within the framework of a theory; and that if the theory then proved to supply a criterion for his self-attribution of ignorance, or error, concerning the status of some particular such report, then—no matter which it might be—he would be willing to reappraise it. But this implication should make us wary. If we deem that the occurrence of no *actual* attempt at theorizing, or only of an unsuccessful attempt, is quite compatible with the existence of all the requisite distinctions, so long as the private linguist acknowledges that the drawing of each of them could be appropriate under theoretical pressure, we would seem committed to the view that his 'observation' statements may already be credited with a meaning which will be more or less invariant across the various theory-inclusive systems of beliefs in which they may come to be embedded. If that commitment seems untroublesome, we shall do well to remember how difficult it has proved to sustain the corresponding claim about the content of reports of observation in the context of our ordinary scientific theories (and, indeed, to sustain so much as the claim that there is a well-defined class of 'observation statements' at all).

There is a second, more fundamental point. A theory ought not to require the truth of any particular observation statement *unconditionally*. Naturally, depending on whether a particular observation statement is supposed to be true or false, a theory may enjoin a differential assessment of other such statements; but *both* the union of the theory with any single observation statement *and* its union with the negation ought to express coherent possibilities. We have not pressed the Fallibilist for an account of the *sources* of possible error in the private linguist's reports of his sensations; but such an account is owing, and it is hard to see how a satisfactory story could avoid sanctioning the possibility that error might take place, in any particular case, *whatever* theory, if any, the subject held about the dependable patterns of occurrence among his sensations. So much is demanded by the analogy with observation and ordinary empirical theorizing which the Blackburn/Harrison proposal imposes. If, accordingly, it is on the intelligibility of error that the supporter of private language wishes to base the needed (a)/(a)*–(b)/(b)* contrast, it is perfectly fair to insist that he explain how the private linguist might achieve a state of information of type (i) compatibly with a reasonable confidence in whatever theory he had arrived at.

To relate the point to the example above. A satisfactory explanation of what it is for the private linguist to have been in error about

'A sensation of type S_2 occurred' ought to represent it as a possibility consistent with the truth of his belief in H_1 and H_2. The conjunction of 'X erroneously believed "A sensation of type S_2 is occurring"' with H_1 and H_2 ought therefore to represent a conceivable, self-consistent item of information—differing, for instance, from the information represented by the result of replacing 'erroneously' by 'erroneously but excusably'. But, so long as we stay with the principle about information content which has governed the discussion to this point—that two items of information differ only if there could be such a thing as having one without the other[17]—no way presents itself for giving sense to that sort of contrast. There is, simply, no prospect of an account of what it would be for X to have just the theoretical information incorporated in H and H_2 plus reason to think that his prior acceptance of 'A sensation of type S_2 is occurring' was mistaken.

Let us see where we have got to. I have argued for two theses. *Thesis A* affirms that the private linguist can satisfactorily draw the (a)/(a)*–(b)/(b)* contrast for a particular S_k only if he can draw that contrast both for *every* such statement which, he conceives, may contribute with S_k to a body of evidence for or against his theory, and for *the negations* of all such statements. The grounds for this thesis are the principle (T), the argument for the 'rotten apple' effect, and the doubt about the capacity of the private linguist to justify the play he will need to make with the notion of consistency unless the thesis is respected. *Thesis B* affirms that if—which is the private linguist's situation—*only* theory can assist in the drawing of these contrasts, then they are drawn satisfactorily only if drawn within the framework of a *single* theory. In particular, it is not good enough merely to advert to the possibility of different theories which might repair the omissions of whatever theory the private linguist happens to respect. The grounds for this thesis are two. First, the private linguist needs to suppose that the 'omitted' cases *already* enjoy the content which they would possess when harnessed to a theory which

[17] Of course, this principle may be called into question; and it will be the task of the next section to review how. But what seemed interesting about the Blackburn/Harrison proposal was exactly that it promised the private linguist the prospect of drawing the requisite sorts of distinction in *operational* terms, compatibly with the requirements of our governing principle. (The invocation of theory would be pointless if that principle were no longer accepted, since the enjoyment of an appropriately contrasting content by instances of (a)* and (b)* would no longer have to answer to the possibility of states of information of type (i)–(iv).)

actually drew the (a)/(a)*–(b)/(b)* contrast for them—otherwise he will run afoul of the 'rotten apple' and consistency points; and he holds out a substantial hostage to fortune if he undertakes to defend that supposition against the familiar difficulties which beset the notion of the continuity of 'observational' contents through theoretical change. Second, it is part of our ordinary concept of observational error that it may occur whatever theory we hold. Accordingly the private linguist has not attained a satisfactory conception of what it is to be in error about, for example, the occurrence of an instance of S_2, if he does not understand what it would be both for that to be so and for his current theory to be correct. The operationalism implicit in the Blackburn/Harrison response—that the 'seems right'/'is right' distinction should be one the private linguist can actually draw—entails, accordingly, that reason to suspect obervational error—error in the appraisal of some S_k or its negation—should always be attainable compatibly, at least when taken in isolation, with *whatever* theory the subject is inclined to endorse. (Or, if it does not, it has suddenly turned suspiciously eclectic.) Of course there is no difficulty in meeting this condition when, as is usual, the criteria for observational error are, at least in part, of a largely non-theoretical sort—discord with others' reports, poor lighting, mislaid spectacles, and so on. But in the present case such criteria *have* to be theoretically generated.

Putting theses A and B together, we arrive at the intriguing result that Blackburn's and Harrison's strategy can be both pointful and successful for the private linguist only if his theory can indeed do for *each* member of the relevant class of 'observation statements' and their negations what the theory in the example does for 'A sensation of type S_3 occurred'. Could there be such a theory, and what would it look like?

Consider the total set of rows of Ts and Fs which specify all possible truth value assignments to the ingredient variables in a formula in n variables of classical sentential logic—let us call this structure of Ts and Fs the *assignment block*. Let us say that one row is an n-transform of another just in case the first results from the second by changing the assignment to exactly n of the variables. Then the theory in the example does what it does for 'A sensation of type S_3 occurred' because the relevant assignment block contains a row—line 5—which falsifies the theory as a whole, assigns T to S_3, has a 1-transform which verifies the theory as a whole and is achieved by changing just the assignment of T to S_3, and has no other

1-transform which verifies the theory as a whole. All that is the formal reflection of the point that the findings which line 5 represents are inconsistent with the theory but can be restored to consistency with it by a reappraisal which is *uniquely* least disturbing of them. The upshot is that if the private linguist wishes to bring what strike him as different sensation types under theoretical rein, and satisfactorily to meet what I am presenting as Wittgenstein's argument in the Blackburn/Harrison way, his 'theory' had better take a form whereby its output for the n relevant types of 'observational statement' corresponds to a classical truth function, ϕ, in n variables with the following features:

For each variable v_i there are rows in the assignment block, $V_j(v_1 \ldots v_n)$ and $V_k(v_1 \ldots v_n)$, such that:

$V_j(v_i) = \mathrm{T}$;	$V_k(v_i) = \mathrm{F}$
V_j falsifies $\phi(v_1 \ldots v_n)$;	V_k falsifies $\phi(v_1 \ldots v_n)$
$V_j{}^*(v_1 \ldots v_n)$, differing from $V_j(v_1 \ldots v_n)$ only in the assignment to v_i, verifies $\phi(v_i \ldots v_n)$;	$V_k{}^*(v_1 \ldots v_n)$, differing from $V_k(v_1 \ldots v_n)$ only in the assignment to v_i, verifies $\phi(v_1 \ldots v_n)$
There is no other 1-transform of $V_j(v_1 \ldots v_n)$ which verifies $\phi(v_1 \ldots v_n)$;	There is no other 1-transform of $V_k(v_1 \ldots v_n)$ which verifies $\phi(v_1 \ldots v_n)$

We now need to do some formal work whose details I relegate to an Appendix. But here are some pertinent results. If $n = 2$, there are no theories of the requisite sort. If $n = 3$ or 4, all the theories of the requisite sort correspond to truth functions which are verified only by a pair of complementary rows in the assignment block; for example, by [TTF] and [FFT]—lines 2 and 7 in the assignment block for H_1 and H_2 above—which would correspond to the theory:

H_3: a sensation of type S_1 occurs if and only if a sensation of type S_2 occurs; and

H_4: a sensation of type S_2 occurs if and only if a sensation of type S_3 does not occur.

Indeed one way to arrive at a suitable theory, irrespective of the size of n, is simply to write up the chain of biconditionals which correspond to any single row in the assignment block in the manner in which $\{H_3, H_4\}$ corresponds to line 2 in the three-variable assignment block tabulated earlier. But this is not the only way of finding a suitable theory if n is greater than or equal to 5. For example, the function in 5 variables which is verified at each of [TTFFT], [FFFTT], and [FTTTF] but nowhere else corresponds to a suitable theory for a language of 5 sensation types; as does that which is verified at each of [TTTTT], [TTFFF], [FFFTT], and [FFTFF], but nowhere else. (The alert reader will spot that the crucial consideration is that each designated row be at least a 3-transform of all the others, and that the designated rows collectively comprise assignments of both T and F to each variable.)

The resulting situation is somewhat peculiar. The feeling that a private language, in particular a private language of sensation, is possible is quite indifferent to the degree of complexity and variety of the material to be described. But if the private linguist happens to have only two sensation types, or if only two out of a larger number of types by which he is afflicted prove to be theoretically tractable, the Blackburn/Harrison strategy is no good to him. And if his inner world is richer, and a threefold or larger variety within it does seem to prove amenable to theoretical regimentation, the chances that his experience will suggest a theory falling within the range of those which deliver all that has been argued to be necessary are signally slim. There are, for example, half as many distinct biconditionally constituted theories, of the sort illustrated, in n types of 'observation statement' as there are rows in the assignment block; so the ratio of the number of such theories to all possible theories, where each possible theory corresponds to a distinct truth function in n variables, is $2^{n-1}/2^{2^n}$. Since the biconditionally constituted theories are the only suitable theories when $n = 3$ or 4, this gives the would-be private linguist a one in 64 chance of meeting Wittgenstein's objection if his theory deals in three sensation types, and a one in 8192 chance if it deals in four! Even the most determined soliloquist ought to be dismayed by such odds. The reader will anticipate that matters get worse as n increases, and so indeed they do. In particular, the limiting

frequency of suitable theories among all possible theories tends to zero as n tends to infinity. (For proof, see the Appendix.)

I conclude that the Blackburn/Harrison objection can at most oblige some refinement of Wittgenstein's claim; it does not, in the end, qualify its force. The argument proposed a necessary condition for a class of statements to qualify as genuinely factual and suggested that private language could not meet it. Not so, said Blackburn, provided the language embeds an appropriate theory. And thus far, except for the case $n=2$, he is entirely right. But we have seen that Wittgenstein's objection can be amplified in such a way that only possession by the private linguist of a very special sort of theory can put him in position satisfactorily to respond to it in the Blackburn/Harrison way. The result is that although we find in favour of the *bare possibility* of private language—or, more accurately, of surmounting the barrier to it adverted to in I.258—enough has been done to undermine the *motivation* of each of the foreseeable species of believer in private language distinguished in section 1. One who believes in the essential privacy of large parts of his mental life will surely want to suppose that his capacity to record its character in terms no one else can have reason to think he understands would be *in no way contingent* on the particular form of the patterns, if any, of concomitance which the various event types display, but would depend only on the adequate functioning of his faculties. One who succumbs to sceptical pressure towards the third-person inscrutability of intention, or to solipsism, will want to say the same. The extra element of contingency might be tolerated if it was associated with high probability; but quite the reverse is true. Blackburn's objection forces a concession which offers absolutely no comfort to Wittgenstein's traditional antagonists.

It remains to draw a further moral. It is a fact that we actually possess an 'operational' grasp of the notions of error and ignorance for statements of every sort, including everything we might incline to regard as a report of observation. One implication of the preceding is that it is wildly unlikely that this could be so unless this grasp owed more to our membership in a language community in which we have faith in others' judgements than to our engagement in theory-building. This is one crucial difference which a community, actual or potential, makes: only by reference to a (potential) community of speakers of his language can a subject guarantee himself

any reasonable likelihood of globally applicable, operational notions of ignorance and error.

9. *Drawing the 'Seems Right'/'Is Right' Distinction Non-operationally*

I warned in section 4 that the interpretation to be recommended would not entirely distance itself from verificationism. Such verificationism as it involves is implicit in the erstwhile governing principle that (at least for a very large class of cases, including relevant instances of (a)/(a)* and (b)/(b)*) statements may be regarded as differing in their informational content only if there could be such a thing as having reason to believe one without the other. For the principle leaves no space for the possibility (as some will see it) of grasping that the informational content of a pair of statements differs although there can be no such thing as having even the weakest reason to accept *either*.

Anyone who was not already dissatisfied with the governing principle when it was introduced but now dislikes its Wittgensteinian consequences should probably have a bad conscience about dramatizing this verificationist aspect. But, in any case, the belief in the possibility that verification-transcendent statements may differ in informational content is not to the purpose for a supporter of private language, since he will not wish to maintain that there is no such thing as having either piece of information conveyed by statements of types (a)/(a)* and (b)/(b)*. Prima facie, then, the verificationist aspect of the governing principle plays no part in the argument.

The reply should be that the principle actually imports verificationist cargo in two places. It entails, as noted, that the informational content of a pair of statements can be contrasted only if there can be reason to believe one of them: but it also entails that if possession of reason to believe either involves reason to believe both, we cannot attain *distinct* conceptions of their informational content, that is, understand them as depicting different states of affairs. And that is tantamount to the assumption that the meaning of a statement cannot be determined by truth conditions over and above the possible grounds for believing it. Since (prescinding from presently irrelevant considerations about conversational practice) grounds for belief and grounds for assertion coincide, the assumption is, in effect, that the contrast between (a)/(a)* and (b)/(b)* has to be drawn by reference

to conditions of warranted assertion; that it cannot coherently be allowed that their assertion conditions might coincide while the *truth conditions* diverged, with the latter divergence supplying the requisite contrast in content.

It would be natural to suppose that this reflection takes the private language issue up into the general dispute between realists and anti-realists in the theory of meaning, which has recently increasingly resembled a dialogue of the deaf. It is true of course that if we were to decide—for example for the sort of reason developed by Dummett—against the realist conception of truth, then that would be that. But I do not think that this is the best way to look at the matter. Rather, suppose we find *against* verificationism in general; that is, we satisfy ourselves that there are at least some areas of discourse for which the sort of idea of truth which the verificationist complains about can be made intelligible and is suitable to serve as the basis of a truth-conditional conception of meaning. Still, the victory over verificationism has to be gained piecemeal: persuading ourselves that the verificationist has no cogent *general* point would still leave open the task, for any particular class of statements where realism is our antecedent conviction, of showing that we have indeed attained the kind of conception of truth which the verificationist believes we cannot attain. Even someone who is persuaded that there are no compelling general arguments for verificationism ought not to find it simply *obvious* that the private linguist can intelligibly draw the needed contrasts in the sort of truth-conditional way proposed. It demands a special idea of truth, which has to be earned.

I continue to concentrate on the example of sensation (though the considerations which follow will generalize). The notion which has to be earned is that the private linguist may simply mistake how he ought to describe an occurrent sensation without its being in any way possible for him, then or later, to acquire evidence of the mistake. One consequence of admitting this idea is that he divests himself of any worthwhile reason for supposing that this situation is not *frequently* realized. When the idea of mistake is bound to (at least a high probability of) detectability, it is possible to get some purchase on the question of how error-prone a particular subject is. But to suppose that any mistake made by the private linguist is going to be undetectable entails that he has absolutely no basis for any view about the frequency of his errors. And that in turn entails that he has no entitlement to regard any of his opinions as likely to be correct.

So one who favours this pre-emptive way with the line of reasoning developed in the preceding section immediately pays a significant price: he surrenders not merely Cartesian authority but even the weaker authority which the Fallibilist is likely to want to claim for his descriptions of those of his mental states which he conceives as private. Ironically enough, he thereby places himself in position to meet the demand which his strategy was designed to avoid: the demand that he give *operational* content to the (a) and (b) contrast. For he now has reason to say that states of information of type (iii) are available, the fact of his believing '*P*' to be true invariably providing absolutely no reason for supposing it to be so. Of course he is no closer than before to explaining how he could perfectly *rationally* enjoy such a state of information, since he has no evident defence against the thought that to allow himself to have any beliefs at all with the relevant sort of subject matter is a kind of passive irrationality. But, in any case, the availability of states of information of type (iii) is no longer to the dialectical purpose: the distinction demanded by Wittgenstein's argument has already been drawn if the private linguist can indeed attain so deeply realist a conception of the truth conditions of his private 'statements'. The question remains whether he can.

There are a number of reasons for doubting it, even if we forgo invocation of the general considerations about concept-acquisition and concept-display on which Dummett, for instance, builds a global anti-realist case. One important line of attack would pivot on whether it is permissible to think of meaning as possessing the strong objectivity which this defence of private language demands: the objectivity implicit in the idea that the meaning of a statement can be settled by a presumably finite set of behavioural and intellectual episodes in such a way as to determine truth value (with appropriate assistance from the world) quite independently of any actual or possible response from those whose understanding might naturally be held to constitute the meaning of the statement. An opposing conception would be that the responses of those who understand it stand to the meaning of a statement rather as someone's behaviour stands to his character: there is conceptual space for a failure of fit—acting out a character, misuse of the statement—but the relationship is somehow, at bottom, still a constitutive one. If anything like this latter conception is right, the meaning of the private linguist's statements could not possibly sustain the utterly fortuitous connections

with his preferred uses of them which the realist defence of private language demands. A strong case can, I believe, be made for this conception without invoking anything which should be regarded as question-begging anti-realist presupposition.[18] But it would take us too far afield to review it here.

There are, however, two rather more specialized lines of thought. Both contend, in different ways, that the realist defence puts the ordinary idea of *intention* under intolerable strain. First, the realist owes an account of the precise mechanism which, quite independently—as we have seen—of the private linguist's beliefs, determines the truth values of his statements. What makes it the case that, say, 'A sensation of type S_2 is occurring' is true, when it is true, independently of the private linguist's view? The answer has to be: resemblance between the occurrent sensation and those of its predecessors whose association with the label 'S_2' established its meaning. But what determines what kind, or degree, of resemblance is *good enough*—qualifies the sensation as a member of the relevant class? Pain, for example, can vary in duration, in intensity, and in a host of other qualitative ways. What is going to settle how far an S_2 sensation may vary, and in what parameters, before it is disqualified? The issue has to turn on the subject's intentions on the baptismal occasions: the resemblance is close enough when it is the kind of resemblance that the subject intended. But the concept of intention has a breadth far in excess of the concept of thought: I need not have explicitly entertained a detailed scenario comprising a specific course of action in specific background circumstances before it can be true to say of me that I intended that that course of action should take place in those circumstances. It is unclear how it would be possible for the concept of intention to have this breadth if we did not take a subject's actual responses as criterial for the character of his former intentions. Suppose now, however, that the private linguist has a sensation which—if you think you understand what it means to say so—is strictly qualitatively unlike any of the baptismal cases, but which he has no hesitation in describing as a further instance of S_2. The realist defence is committed, as we have seen, to the view that there can be no better reason for regarding this response as correct than as incorrect—committed, therefore, to surrender of the ordinary criterial connection just noted. So what is to be regarded as establishing the *content* of the intention which settles whether or not this particular response is correct?

[18] Cf. my 'Rule-following, Meaning and Constructivism', in C. Travis (ed.), *Meaning and Interpretation* (Blackwell, Oxford, 1986).

It is very unclear that the materials are to hand for a satisfactory reply. Suppose the private linguist did not envisage this precise case (again, assume we know what it is to envisage a *precise* case). Then, when we are forced to regard the character of his present response as irrelevant, it is baffling what could give the intention the requisite determinacy, could make it true, or not, that the occurrent sensation should be described as an instance of S_2. Suppose instead, more conveniently, that the private linguist did envisage the precise case. Then the question is whether that can be true indefinitely often; whether, in other words, his intention could have had the *general* content necessary in order to establish a meaning, that is, a rule. Is it possible to envisage indefinitely many 'precise cases'? The requisite rule would have taken a form, I suppose, something like:

> Whenever a situation occurs exactly like this one, 'A sensation of type S_2 is occurring' is true.

But that is open to the simple-minded reproach that a situation 'exactly like this one' is never going to occur again: something is bound to be different. The rule needs some notion of *relevant* similarity. But now the class of 'relevantly similar' situations has got to be indefinitely various, since its members may vary indefinitely in the putatively irrelevant respects. So the question recurs: what can make it true that each of these was intended to count as a case of relevant similarity when all that is available to fix the content of the original intention is the explicit thoughts which the subject entertained at the time, and when only a proper subclass of these situations can have been so explicitly thought about?

The foregoing is, of course, strongly reminiscent of Kripke's Sceptical Argument. But it would be wrong to conclude that it presents a real difficulty for the would-be private linguist only if Kripke's sceptic argues cogently. For there is a difference. Kripke's sceptic gulls his interlocutor into accepting that the content of his previous intentions has to be recoverable—if it is determinate at all—from his previous thoughts and behaviour. The possibility is thereby passed over that the subject be granted non-inferential recall of previous

intentions, that it be analytic of the concept of intention that people are, by and large, authoritative about their own intentions past and present. The *special* problem to which the realist defence of private language is vulnerable is that, by transcendentalizing the notion of mistake and hence that of the subject's reliability, it abrogates the right to call on this aspect of the concept of intention and thereby—I suggest—leaves Kripke's sceptic with a clear run on goal.

The second way in which the realist defence places the concept of intention under strain is perhaps more immediate. Meaning, it is again platitudinous to say, is normative: it is because statements have meaning that there is such a thing as correct, or incorrect, use of them. This is a platitude which the truth-conditional conception of meaning, like any other, must respect. The story will be, presumably, that the making of a statement is basically correct when its truth conditions are realized (though it may be criticizable on all sorts of other grounds); and is basically incorrect when its truth conditions are not realized (though it may then be excusable on all sorts of grounds). But this story respects the normativity of meaning only if the making of statements just when their truth conditions are realized constitutes a feasible policy. And it is important to realize that its feasibility is not guaranteed if one merely accepts—which of course an anti-realist would not in certain controversial cases—that there is such a thing as the truth conditions of the statement being realized or not. It is necessary, additionally, that there can be such a thing as *aiming* at making or assenting to a statement only when its truth conditions are realized. It is about that that there is a doubt in the present case.

The trouble, once again, flows from the consequence of the realist's defence that the private linguist's believing 'A sensation of type S_2 is occurring' is no longer even the weakest ground for supposing it true. The truth of an 'observation statement' in the private language has become utterly dissociated from any practical criterion. Can the private linguist still aim at making such statements only when they are true—is there any such thing as having that intention? The question, in effect, is whether there can be such a thing as aiming at a transcendent target: a target such that there is no criterion for saying of any particular shot whether it hits, or is likely to hit, or not, and hitting or missing which can have no consequences for the course of your own or another's future experience.

To sample the flavour of the difficulty, try the following thought

experiment. Suppose you are confronted with a pair of sealed boxes, one empty and the other containing an Egyptian scarab. The scarab will vaporize instantaneously and without trace if the seal on its box is disturbed. And there is no exterior sign—no difference in weight, or rattle, and so on—to betray which box is which.[19] Suppose finally that the craftsman who originally made and sealed the boxes is dead and has left no record to help you, although it is quite certain that one and only one of the boxes does contain a scarab. And now suppose you are invited to pick the box with the scarab in by placing your finger on top of it. Can you so much as *try* to do so? What exactly will distinguish your performance from a response to the invitation to pick the empty box, or just to pick a box? Not, at any rate, the exact form taken by your behaviour; that will be consistent with its being a response to any one of the three invitations. Nor the fact that it is a response to one in particular of the invitations, since—if each of the three aims really is possible—it ought to be possible, out of contrariness, to aim, say, at the scarab when invited to pick the empty box. So the distinguishing feature or features have to be *interior*—what are they?

If the circumstances were different, the distinction could emerge in your response to the discovery that your chosen box did, or did not, contain the scarab—disappointment, self-satisfaction, relief, indifference, and so on. So could it not, in the circumstances described, be true of you when you make your choice that if it *were* possible to uncover the scarab and if it *were* in your chosen box, then your response *would* be one of relief, for example? Perhaps. But making the sought-for difference reside in the truth or falsity of such counterfactual conditionals doesn't make it any clearer how it is supposed to be apparent *to you* which intention you have; the problem merely shifts to that of explaining how you can tell which among such counterfactuals are true of you. And there would, in any case, be a strong suspicion of circularity about an attempted counterfactual analysis of the relevant differences, since it would have to be stipulated that the counterfactual circumstances be ones in which your intentions had *remained the same*. But most serious of all: the obstacles for the detection of the beetle are at best causal; whereas the realist defence of private language makes it a *conceptual* truth that there is no detecting the truth value of the private linguist's 'observation statements'.

[19] The empty box is lined with a quantity of the same volatile material sufficient to equalize the weights.

Once that is recognized, it ought to seem quite unclear what is being said if it is claimed to be true of the private linguist for instance that if he *were* to get an independent check on the truth value of 'a sensation of type S_2 is occurring', he would be gratified, or whatever.

10. Conclusion

The net effect of the foregoing is to confront a proponent of the possibility of private language with a dilemma. One way or another he has to draw the 'seems right'/'is right' contrast. If he accepts the need to do so operationally, by reference to practical criteria, our finding was that he will be able to do so in a satisfactory way only in very special, at best unlikely circumstances; and that an element of contingency will thereby intrude into his central claim which is quite foreign to his original conviction. If, on the other hand, he claims the distinction need not be operational but may be drawn purely truth-conditionally, he commits himself to the unwelcome consequence that the private linguist can have no satisfactory basis for his belief in his general competence to practise the private language; and this, besides involving a substantial hostage to fortune in his commitment to a highly objective notion of meaning, generates special difficulties in explaining what constitutes the truth of a statement in the private language and what constitutes the private linguist's aiming at the truth. The answer which all this suggests to the title question is therefore: 'Probably'. The ball, anyway, is in the opposition's court.

APPENDIX*

On the number of classical truth functions in n variables which are congenial to the Blackburn/Harrison purpose.

I

The assignment block (cf. p. 247) consists of 2^n distinct n-fold rows of Ts and Fs. Let us say that one such row is an m-transform of another just in case they differ in exactly m places, $1 \leqslant m \leqslant n$. A truth function, ϕ, in n variables

* This is not an exclusive for formal logicians. At least, I have *tried* to present the reasoning which follows in a way which should be intelligible to a logical, rather than logically expert, reader with a memory of school arithmetic that can at least be jogged, and a rudimentary grasp of elementary logic and set theory.

may be identified with that subset of these rows for which it yields the value T. Call the members of such a subset, ϕ, the ϕ-selected rows. Our question is this:

> How many subsets are there, as a function of n, with the feature that for each variable there is *both* a ϕ-unselected row where that variable is assigned T, which has a ϕ-selected 1-transform involving change in just that assignment, and which has no other ϕ-selected 1-transform, *and* a ϕ-unselected row where that variable is assigned F, which likewise has a ϕ-selected 1-transform involving change in just that assignment, and which has no other ϕ-selected 1-transform?

I can at present provide no general answer to this question; nor even a proof that the ratio of the number of such *congenial* subsets over 2^{2^n}—the number of all possible classical truth functions (theories) in n variables—is strictly decreasing as n increases. But, where $\lambda(n)$ is the number of such congenial subsets, we can justify the following claims:

(a) for $n=2$, $\lambda(n)=0$;

(b) for $n=3$, $\dfrac{\lambda(n)}{2^{2^n}}=\dfrac{1}{64}$;

(c) for $n=4$, $\dfrac{\lambda(n)}{2^{2^n}}=\dfrac{1}{8192}$;

(d) for n in general, $\lim\limits_{n\to\infty}\dfrac{\lambda(n)}{2^{2^n}}=0$.

I conjecture that $\dfrac{\lambda(n)}{2^{2^n}}$ *is* strictly decreasing as n increases. And I suggest that the *philosophical* import of Claim (d) is the same as that of the truth of this conjecture.

II: Claims (a), (b), and (c)

It will be apparent that any congenial subset will contain more than one row, selected from opposite halves of the assignment block (given the usual conventions for listing the Ts and Fs). More generally, the selected rows must collectively involve assignments of both T and F to each of the n variables. Each selected row is associated with n 1-transforms. Let us say that an unselected row is *useful* with respect to a variable or the negation of a variable, just in case it is a 1-transform of a selected row in which that variable is assigned F/T and of no other selected row. So: we have to select rows in such a way as to generate at least $2n$ useful unselected rows, each useful in a different respect. (I shall say that a selected row gives a *decision* with respect to a variable or its negation just in case some unselected 1-transform of it is useful with respect to that variable or its negation.)

Case (a), $n=2$:
ϕ must generate at least $2n$, $=4$, useful unselected rows. But the assignment block contains only 4 rows. So there is no such ϕ; whence Claim (a).

Case (b), $n=3$:
ϕ must generate at least $2n$, $=6$, useful unselected rows. The assignment block contains only 8 rows. Hence ϕ must contain at least and at most 2 rows, which—since they must together involve assignments of both T and F to each of the 3 variables—must be *complementary* (where rows a_i and a_j in the n-variable assignment block are complementary just in case they are n-transforms).

So

$$\frac{\lambda(3)}{2^{2^3}}=\frac{2^{3-1}}{2^{2^3}}=\frac{1}{64}.$$

Case (c), $n=4$:
Note the following three points. First, each ϕ-selected row is potentially associated with 4 useful rows. This potential is frustrated, however, if ϕ contains rows which are 2-transforms of each other, for then some of their 1-transforms will coincide, so will not be useful. If, for example, ϕ contains [TFTT] and [TTFT], we lose the potentially useful rows [TFFT] and [TTTT] since they are each 1-transforms of both those rows. More generally, a pair of selected 2-transforms can be associated with only 4, rather than 8, useful rows; and these 4 can be useful in only 2 respects—[FFTT] and [FTFT], for example, are both still useful when ϕ is as above; but they are useful in the *same* respect, namely with respect to the negation of the first variable.

Second, there is no point in ϕ's containing rows which are 1-transforms of each other since the useful rows associated with them will be useful in the same three respects—determined by the asignments of Ts and Fs shared by the selected rows—and no useful rows will be associated with them with respect to the variable over whose assignment they differ, nor with respect to its negation. So: if ϕ is congenial and contains a pair of 1-transforms, there is a congenial ϕ which omits one member of the pair. *Lemma*: if there are no k-fold congenial sets not involving 1-transforms, there are no $k+1$-fold congenial sets which do involve 1-transforms. So we can, for the moment, disregard the possibility of selecting 1-transforms.

Third, if ϕ contains (at least) three rows, no two of which are 1-transforms, then it must contain at least one pair of 2-transforms.*

* *Proof*: let a_i, a_j, a_k be three distinct rows no two of which are 1-transforms. Suppose, for reductio, that no two are 2-transforms either. So a_i is a 3- or 4-transform of a_j, and a_j is a 3- or 4-transform of a_k. If a_i is a 4-transform of a_j, then $a_i=a_k$, contrary to hypothesis, if a_k is a 4-transform of a_j; and a_i is a 1-transform of a_k, contrary to hypothesis, if a_k is a 3-transform of a_j. So a_i must be a 3-transform of a_j. Likewise a_j must be a 3-transform of a_k. But then, if $a_i \neq a_k$, they must be 2-transforms, since a_k must differ from a_j both in that assignment which a_j and a_i have in common and in two over which they differ.

Consequences: (i) No ϕ, not involving 1-transforms, containing exactly *three* rows can be congenial. *Proof*: since at least two must be 2-transforms, their associated useful rows can be so in only the same two respects, as noted above. The third row can generate at most four more associated useful rows—so we finish with only, at most, six decisions out of the eight we need.

(ii) No ϕ, not involving 1-transforms, containing exactly *four* rows can be congenial. While a fully explicit proof would be a bit involved, the reason, basically, is that any such ϕ includes four 3-fold subsets, each of which must contain at least one pair of 2-transforms. There are then two possibilities. *Either* ϕ includes two *disjoint* pairs of 2-transforms—in which case we get at most four decisions out of the eight we need. *Or* ϕ includes a *transitive triple* of 2-transforms—a_i a 2-transform of a_j; a_j a 2-transform of a_k; and a_i a 2-transform of a_k—which, as the reader may verify, can collectively contribute only a *single* decision so that, given that the fourth selection can contribute at most a further four, ϕ can yield at most five decisions out of the eight we need.

I leave it to the reader to verify that as ϕ increases to five, six, seven, or eight rows, the proliferation of disjoint pairs, and/or of transitive triples, of 2-transforms invariably results in fewer than the requisite eight decisions. As with $n=3$, the only congenial sets for $n=4$ are pairs of complementary rows. So

$$\frac{\lambda(4)}{2^{2^4}}=\frac{2^{4-1}}{2^{2^4}}=\frac{1}{2^{13}}=\frac{1}{8192}.$$

It is not true, however, even for $n=5$, that every congenial ϕ consists in a pair of complementary rows. For example, [TTFFT, FFFTT, FTTTF] is a congenial *three*fold ϕ and [TTTTT , TTFFF , FFFTT , FFTFF] a congenial *four*-fold ϕ (each selected row being at least a 3-transform of all the others). So the natural conjecture, that

$$\frac{\lambda(n)}{2^{2^n}}=\frac{1}{2^{(2^n-n)+1}}$$

is false.

III: Claim (d)

Strategy of proof:
First some further definitions. Let $\tau_i a$ be the 1-transform of row a which differs from it in the ith place. (Let $\tau_o a = a$.) Let an r-tuple $[a_1, \ldots, a_r]$ of rows

from the assignment block be *F-good*, respectively *T-good*, iff the *i*th place of a_i is F, respectively T. Let such an *r*-tuple *fit* a particular subset (truth function) ϕ iff, for each a_i, $\tau_i a_i \varepsilon \phi$ but neither a_i itself nor any other 1-transform of $a_i \varepsilon \phi$.

Plainly, where n is the number of variables with which we are concerned, ϕ is congenial iff it is fitted both by a T-good n-tuple and by an F-good n-tuple. (For the constituents of the two n-tuples will thereby comprise $2n$ useful rows, each useful in a different respect.)

Let S be the set of all rows in the 2^n-fold assignment block. And let $Y(a_1, \ldots, a_r)$ be the set consisting of some $a_1, \ldots, a_r, \varepsilon S$, and all their 1-transforms. Let an r-tuple $[a_1, \ldots, a_r]$ be *independent* if $a_1, \ldots, a_r$ are all distinct and none is a 1- or 2-transform of any other. And let an $(r+1)$-tuple $[a_1, \ldots, a_r, a_{r+1}]$ be *barely dependent* if $[a_1, \ldots, a_r]$ is independent but a_{r+1} is a 1- or 2-transform of at least one of $a_1, \ldots, a_r$.

Finally let $m = \frac{n}{2}$ if n is even, $\frac{n+1}{2}$ if n is odd. Let C be the set of subsets, ϕ, of S such that some independent F-good m-tuple fits ϕ. And, for each $r \geqslant 1$, $\leqslant m$, let D_r be the set of subsets, ϕ, of S such that some barely dependent F-good $(r+1)$-tuple fits ϕ.

The proof of claim (d) hinges on the observation that if ϕ is congenial, then $\phi \varepsilon \{C \cup D_1 \cup D_2 \cup .. \cup D_{m-1}\}$.

Proof: if ϕ is congenial, it is fitted by an F-good n-tuple $[a_1, \ldots, a_n]$. Let $[a_1, \ldots, a_m]$ be the first m terms of $[a_1, \ldots, a_n]$. If $[a_1, \ldots, a_m]$ is independent, then $\phi \varepsilon C$. If $[a_1, \ldots, a_m]$ is not independent, let $[a_1, \ldots, a_i]$ be its longest independent initial segment (which might just be a_1). In that case $[a_1, \ldots, a_i, a_{i+1}]$ is a barely dependent F-good $(i+1)$-tuple which fits ϕ; so $\phi \varepsilon D_i$.

Accordingly the number, $\lambda(n)$, of congenial such subsets, ϕ, is less than or equal to the cardinality of the union of $C, D_1, D_2, \ldots, D_{m-1}$; which is in turn less than or equal to the sum of the cardinalities of $C, D_1, D_2, \ldots, D_{m-1}$. The proof now proceeds by estimation of upper bounds on the cardinality of these sets. We establish

Lemma A:

$\overline{\overline{C}} \leqslant 2^{2^n - n}$, for arbitrary n, and

Lemma B:

$$\overline{\overline{D_1}} + \overline{\overline{D_2}} + \ldots . \overline{\overline{D_{m-1}}} \leqslant 2^{2^n - \left(\frac{n}{2}+1\right)},$$

provided that $n \geqslant 32$.

It follows that, for $n \geqslant 32$,

$$\lambda(n) \leqslant 2^{2^n - n} + 2^{2^n - \left(\frac{n}{2}+1\right)}.$$

Since, for $n > 2$, $2^n - n < 2^n - \left(\frac{n}{2}+1\right)$, we may infer that

$$\lambda(n) < 2 \times 2^{2^n - \left(\frac{n}{2}+1\right)}, = 2^{2^n - n/2}, = 2^{2^n} \times 2^{-n/2}, = 2^{2^n} \times \frac{1}{2^{n/2}}, = \frac{2^{2^n}}{2^{n/2}}.$$

Hence, dividing each side by 2^{2^n}, we have

$$\frac{\lambda(n)}{2^{2^n}} < \frac{1}{2^{n/2}}, \text{ for } n \geqslant 32.$$

Whence, since $\lim_{n \to \infty} \frac{1}{2^{n/2}} = 0$, Claim (d) follows.

Let us therefore establish the two lemmas.

Proof of Lemma A

We prove three sublemmas:

Lemma A1: if $a_1, \ldots, a_r$ are any rows in S, the number of subsets ϕ such that $[a_1, \ldots, a_r]$ fits ϕ is less than or equal to $2^{\overline{\overline{S - Y(a_1, \ldots, a_r)}}}$, that is, 2 to the power of the number of rows in S which are distinct from $a_1, \ldots, a_r$ and are not 1-transforms of them; which is the number of sets of such rows.

Proof. If $[a_1, \ldots, a_r]$ fits ϕ, then the only members of $Y(a_1, \ldots, a_r)$ which are members of ϕ are each $\tau_i a_i$, $1 \leqslant i \leqslant r$. Hence if $[a_1, \ldots, a_r]$ fits distinct ϕ and ϕ', then the membership of ϕ and ϕ' must be the same as far as $a_1, \ldots, a_r$ and their 1-transforms are concerned. Hence ϕ and ϕ' must differ by virtue of one, or both, containing k-transforms, $k \geqslant 2$, of some of $a_1, \ldots, a_r$ which the other does not. So the number of distinct such sets cannot be greater than the number of sets of such k-transforms of $a_1, \ldots, a_r$, $= 2^{\overline{\overline{S - Y(a_1, \ldots, a_r)}}}$.

Lemma A2: if $[a_1, \ldots, a_r]$ is independent, then

$$Y(a_1, \ldots, a_r) = r \times (n+1).$$

Proof. Clearly $Y(a_p) = n+1$; that is the set consisting of a_p and all its 1-transforms is an $(n+1)$-fold set. So it suffices to reflect that, since each element of $a_1, \ldots, a_r$ is at least a 3-transform of every other, they have no 1-transforms in common.

Lemma A3: for each r, there are at most $2^{r \times (n-1)}$ independent F-good r-tuples.

Proof. For each i, there are 2^{n-1} rows with F in the ith place. Hence there are $(2^{n-1} \times 2^{n-1} \times 2^{n-1} \times \ldots r$ times) ways of selecting an F-good r-tuple, so $2^{r \times (n-1)}$ F-good r-tuples altogether.

By Lemma *A*1, the number of subsets, ϕ, such that a particular (independent) $[a_1, \ldots, a_m]$ fits ϕ is less than or equal to

$$2^{2^n - \overline{\overline{Y(a_1, \ldots, a_m)}}}.$$

By Lemma *A*2, if $[a_1, \ldots, a_m]$ is independent, then

$$\overline{\overline{Y(a_1, \ldots, a_m)}} = m \times (n+1).$$

So the number of subsets fitted by a particular independent $[a_1, \ldots, a_m]$ is less than or equal to

$$2^{2^n - (m \times (n+1))}.$$

Since, by Lemma *A*3, there are at most $2^{m \times (n-1)}$ independent F-good m-tuples, the number of subsets fitted by some independent F-good m-tuple or other, $= \overline{\overline{C}}$, can be at most

$$2^{2^n - (m \times (n+1))} \times 2^{m \times (n-1)},$$
$$= 2^{2^n - mn - m + mn - m} = 2^{2^n - 2m}.$$

Since $2m \geqslant n$,

$$\overline{\overline{C}} \leqslant 2^{2^n - n}, \text{ Q.E.D.}$$

Proof of Lemma B

We prove two further sublemmas:

*Lemma B*1: if $[a_1, \ldots, a_{r+1}]$ is barely dependent, then

$$\overline{\overline{Y(a_1, \ldots, a_{r+1})}} \geqslant (r \times (n+1)) + (n+1-2r).$$

Proof. By Lemma *A*2, $\overline{\overline{Y(a_1, \ldots, a_r)}} = r \times (n+1)$. And $\overline{\overline{Y(a_{r+1})}}$, of course, is $n+1$. So Lemma *B*1 holds provided $Y(a_{r+1})$ does not contain *more* than $2r$ members which are also members of $Y(a_1, \ldots, a_r)$. Suppose, for reductio, that there are more than $2r$ such. Then (*) there must be some a_p in $[a_1, \ldots, a_r]$ such that $Y(a_p)$ and $Y(a_{r+1})$ have at least three members in common. Now a_p and a_{r+1} cannot be k-transforms, for any $k \geqslant 3$, or $Y(a_p) \cap Y(a_{r+1})$ would be empty. But they are distinct, by the definition of 'barely dependent'. Hence there are two cases:

Case (a): a_p and a_{r+1} are 1-transforms. Then $Y(a_p) \cap Y(a_{r+1}) = \{a_p, a_{r+1}\}$, contrary to (*).

Case (b): a_p and a_{r+1} are 2-transforms, differing in the ith and jth places. Then $Y(a_p) \cap Y(a_{r+1}) = \{\tau_i a_p, \tau_j a_p\}$, again contrary to (*).

*Lemma B*2: for each r there are at most

$\frac{r \times n \times (n+1)}{2} \times 2^{r \times (n-1)}$ barely dependent F-good $(r+1)$-tuples.

Proof. Each such $(r+1)$-tuple is constituted by an initial independent F-good r-tuple, plus a 1- or 2-transform of one of the first r elements. By

Lemma *A*3, those first r elements may be chosen in at most $2^{r \times (n-1)}$ ways. Now, any $a_s \in S$, has n 1-transforms and $(n-1+n-2+ \ldots +1)$ 2-transforms; so $\frac{(n+1) \times n}{2}$ 1- and 2-transforms. Hence, given that the first r elements have been chosen, the $r+1^{st}$ may be chosen in at most

$$\frac{r \times n \times (n+1)}{2} \text{ ways.}$$

(Of course, some of these choices will not be consistent with F-goodness. But remember that we are estimating an *upper* bound.) Thus the total number of ways of choosing the first r, plus the $r+1^{st}$ element, will not exceed

$$\frac{r \times n \times (n+1)}{2} \times 2^{r \times (n-1)}.$$

Now $\overline{\overline{D_r}}$, = the number of subsets, ϕ, of S which are fitted by some barely dependent F-good $(r+1)$-tuple, will not exceed the product of the maximum number of subsets fitted by a particular $(r+1)$-tuple and the maximum number of barely dependent F-good $(r+1)$-tuples.

By Lemma *A*1 and *B*1, the first is less than or equal to $2^{2^n - ((r \times (n+1)) + (n+1-2r))}$.

By Lemma *B*2, the second is $\frac{r \times n \times (n+1)}{2} \times 2^{r \times (n-1)}$.

So $\overline{\overline{D_r}} \leqslant \frac{r \times n \times (n+1)}{2} \times 2^{r \times (n-1)} \times 2^{2^n - ((r \times (n+1)) + (n+1-2r))}$

which is

$$\frac{r \times n \times (n+1)}{2} \times 2^{rn - r + 2^n - ((rn+r) + (n+1-2r))},$$

i.e. $\quad " \quad \times 2^{rn - r + 2^n - rn - r - n - 1 + 2r},$

i.e. $\quad " \quad \times 2^{2^n - (n+1)}.$

Hence $\overline{\overline{D_1}} + \ldots + \overline{\overline{D_{m-1}}} \leqslant$

$$\frac{n \times (n+1)}{2} \times 2^{2^n - (n+1)} \times [1+2+ \ldots +m-1],$$

$$= \frac{n \times (n+1)}{2} \times 2^{2^n - (n+1)} \times \frac{(m-1) \times m}{2},$$

$$= \frac{n \times (n+1) \times (m-1) \times m}{4} \times 2^{2^n - (n+1)}.$$

Let $\alpha = \frac{n \times (n+1) \times (m-1) \times m}{4}$,

and $\beta = 2^n - (n+1)$,

and $k = \overline{\overline{D_1}} + \ldots + \overline{\overline{D_{r-1}}}$.

So we have shown that $k \leqslant \alpha \times 2^{\beta}$.

Now an interesting arithmetical fact: for values of $n \geqslant 32$, and only for such values,

$$\alpha \times \frac{1}{2^{n/2}} \leqslant 1.$$

So $\alpha \leqslant 2^{n/2}$.

So $k \leqslant 2^{n/2} \times 2^{\beta}$, $= 2^{\beta + n/2}$, $= 2^{2^n - (n+1) + n/2}$,

$= 2^{2^n - (n/2+1)}$, which is Lemma *B*.

Q.E.D.

CHAPTER 9

TRANSCENDENTAL ANTHROPOLOGY*

JONATHAN LEAR

1. Even a sympathetic reader of Wittgenstein's later philosophy must, I think, conclude that it represents an unfinished work. The *Philosophical Investigations*, as Wittgenstein himself says, 'is really only an album': discrete paragraphs are juxtaposed and the reader is left to extract an argument, or a point of view. This was not the result of a deliberate attempt at aphorism—to be a latter day Heraclitus—but of what Wittgenstein called failure 'to weld my results together' into a continuous whole.[1] Biographically speaking, there may be many reasons for this failure. But from a philosophical point of view the interesting question is whether this failure had to occur because Wittgenstein was pursuing disparate strands of thought which cannot coherently be reconciled. The two strands in greatest conflict are what I shall call his *anthropological* and his *transcendental* approaches to philosophy.

Language, for Wittgenstein, is one of the many activities in which men engage; and if it is to be understood, it must be seen as embedded in the context of men living their lives. A 'language-game' is not merely a language, but a 'whole, consisting of the language *and the actions into which it is woven*'.[2] And 'the term "language-game" is meant to bring into prominence the fact that the *speaking* of a language is part of an activity, or of a form of life.'[3] This would suggest that the proper study of language requires that one take an

 I would like to thank the National Endowment for the Humanities for a Fellowship for Independent Study and Research, during which this paper was written. A draft of this paper was delivered to the Institut für Philosophie, Freie Universität, Berlin, and I would like to thank the participants for their comments. John McDowell, Philip Pettit, Timothy Smiley, and Bernard Williams read a draft of this paper and offered valuable comments.

[1] Ludwig Wittgenstein, *Philosophical Investigations* (hereafter *PI*) (Blackwell, Oxford, 1968), p. vii.

[2] Ibid. I.7.

[3] Ibid. I.23.