On Being Forced to a Conclusion

Jonathan Bennett

from: *Proceedings of the Aristotelian Society* suppl. vol. 35 (1961), pp. 15–34.

1. Introduction

The only way to settle conclusively what any part of a language means is to discover the circumstances, both linguistic and non-linguistic, in which the speakers of the language are prepared to use it. This is not a new doctrine, but Wittgenstein gave it new life by dramatising the following question: If someone used an expression in a radically non-standard way, could anything he said about his state of mind convince us that he nevertheless meant it in a standard way? To answer 'No' to this, and to generalize that answer, is to say that the last-resort criteria for what something means lie in the way in which it is used—a fairly plain statement which I shall call 'the behavioural theory (of meaning)' and with which I shall here have no quarrel at all. What I wish to do in the following pages is to consider the relationship between the behavioural theory and some aspects of the concept of proof.

It is beyond dispute that one can be led, by one's acceptance of certain premisses, to accept a certain conclusion. There is, though, a problem about the nature of this 'leading'. On the one hand, it is usual to think of it as sometimes having the nature of a forcing: 'If you say that, you are *committed* to admitting this also; you cannot accept the one and reject the other.' On the other hand, the behavioural theory of meaning makes it difficult to see how there can possibly be such a relationship between the premisses and the conclusion of any proof.

There are many ways of bringing out the apparent clash between the behavioural theory and the notion of logical forcing or committal. Perhaps the clearest of them arises from asking how there can be room for a concept of committal in a purely behavioural study. In this spirit, we might grant that our knowledge of some of the ways in which the parts of a language are used may lead us to expect to find certain sorts of further use and not others; and we could compare such expectations with those of an anthropologist who finds that a certain society has a matriarchal system of authority and then proceeds to investigate its inheritance laws. If he finds that these are patrilinear, he may well be surprised, but he will not say that this aspect of the society's behaviour is made wrong by, or is inconsistent with, other aspects. He would indeed be puzzled to know what could be meant by the suggestion that a society's behaving in certain ways night make it in some way incorrect for it to behave in certain other ways.

He will of course grant that an individual person may be committed by some of his actions, through laws accepted within his society, to behave in certain other ways; and there

is no problem about what is meant by saying that something a person actually does clashes with other things he does, by the rules of his society. Similarly, there is no problem about what is meant by saying that some individual has committed an elementary inferential howler: he has committed a howler if most people would say that it was a howler. But there is a problem about saying, what we should normally want to say, that a move is universally declared to be a howler because it *is* a howler—that if everybody committed it then it would simply be a logical mistake which everybody made. To say this is to parallel the absurd remark that two aspects of a society's behaviour may clash with one another.

It is necessary to insist that our concern with any language is a concern with behaviour, with the noises which people make, the marks they put on paper, and so on. It sounds queer, no doubt, to speak of someone as being debarred from making a certain noise by his prior utterance of certain other noises; it would be more usual to say that he is debarred from assenting to a certain proposition by his prior assent to certain other propositions. But we can cheerfully grant that the meanings create the conflict while still insisting that the noises, and the circumstances of their utterance, are the whole story; for, if the behavioural theory is correct, then to mean such-and-such by a noise is just to be disposed to use it in this-and-that other way as well. It is clear, after all, that our only evidence for what someone means by an expression is how he is prepared to use it. It could be claimed that the uses constitute only evidence, that they are merely pointers to the real meaning which in some way underlies them. In this paper I am taking my stand on the belief that this is a mistake which has been definitively exploded by Wittgenstein, and that use can therefore be taken not just as evidence for, but also as constitutive of, meaning.

Especially be it noted that to lay down rules governing the use of noises is itself only to utter yet more noises. The stress on rules, which is legitimate in itself, could mislead us into denying that logical committal is reducible to relations amongst complex sets of noises. Wittgenstein saw this danger, and insisted at length that problems about meanings cannot be settled just by an appeal to rules, because there will always remain the problem of the meanings of the rules. Kant made the same point less archly, in a charming passage in the *Critique of Pure Reason* (A 133 = B 172):

> To give general instructions how we are to subsume under rules. . . could be done only by means of another rule. This in turn, just because it is a rule, again demands guidance from the judgment. And so we see that, while the understanding is capable of being instructed and equipped with rules, judgment is a peculiar talent which cannot be taught, only practised . . . For although an abundance of rules borrowed from the insight of others may indeed be proffered to—as it were, grafted on to—a limited understanding, the power of rightly employing them must belong to the learner himself; and in the absence of such a natural gift no rule that may be prescribed to him for this purpose can ensure against misuse.

In short: the uses of words cannot be laid down in a fool-proof fashion in a rule, because the rule itself is a verbal item whose own use is vulnerable to fools, in just the same way if not to the same degree as are the word-uses which it regulates.

2. 'The philosophical problem of necessity'

I have referred to a difficulty about saying, what we usually do say, that a proof makes explicit something to which we were committed before the proof was ever constructed. If

the difficulty is insuperable, and if we therefore cannot allow that 'the mathematical proof drives us along willy-nilly until we arrive at the theorem', then we can hardly avoid saying positively that 'at each step we are free to choose to accept or reject the proof' and that it is this element of choice which has wrongly been omitted from the traditional account of proof.

The quoted expressions are from Michael Dummett's 'Wittgenstein on Mathematics' (*Philosophical Review*, 1959), a lucid and eloquent account of the clash between the committal-concept and the choice-concept of proof. Dummett favours the committal-concept, but cannot see how to make it literal and specific. I hope to lay the foundations for doing just that, but, for a reason which will emerge, I wish first to explain the relationship between Dummett's paper and what I am trying to do here. My central concern is not with the fact that Wittgenstein favoured the adoption of a choice-concept of proof, but rather with what I think to be the only viable (not valid) basis for such an adoption, namely a rigorous application of the behavioural theory of meaning. A good deal of this basis can be found in Wittgenstein's writings, but Dummett seems disinclined to explore it. 'Such considerations'—he says on p. 331—'seem to belong to the theory of meaning in general rather than having any particular relevance to the philosophy of mathematics.' But what we have here *is* a problem in 'the theory of meaning in general', and in the philosophy of mathematics only as a special case; and there is fairly conclusive evidence that Wittgenstein too saw the problem in this way.

Now, if we dislike the emphasis on choice but accept the theory of meaning from which it has arisen, our task is to try to show that the link between the two can after all be broken because the concept of logical committal is after all consistent with the behavioural theory. But Dummett does not see the task in this way, since, as I have implied, he is not asking how we can avoid the implausibilities in the choice-concept account of proof while doing justice to the behavioural theory of meaning. He is asking how, confronted as we are by 'the philosophical problem of necessity', we can find a more plausible solution of it than Wittgenstein's:

> The philosophical problem of necessity is two-fold: what is its source, and how do we recognise it? God can ordain that something shall hold good of the actual world, but how can even God ordain that something is to hold good of all possible worlds? We know what it is to set about finding out if something is true; but what account can we give of the process of discovering whether it must be true? (p. 327).

I want to make as clear as I can the relationship between these two questions and the one which I am asking.

It is clear that Wittgenstein (the bluff, outspoken, unequivocal Wittgenstein of Dummett's paper) has an answer to both Dummett's questions: the source of necessity is convention, and necessity is recognised through the examination of conventions. But these answers are possible only because 'for him the logical necessity of any statement is always the *direct* expression of a linguistic convention. That a given statement is necessary consists always in our having expressly decided to treat that very statement as unassailable' (p. 329), and this immediately banishes the concept of logical committal in favour of the choice-concept of proof. The currently popular alternative to this 'full-blooded conventionalism' says Dummett, is a 'modified conventionalism' according to which 'some necessary statements are straightforwardly registers of conventions we have laid down; others are more or less remote consequences of conventions' (p. 328). He is dissatisfied with this modified conventionalism. . .

> . . . since it leaves unexplained the status of the assertion that certain conventions have certain consequences. It appears that if we adopt the conventions registered by the axioms together with those registered by the principles of inference, then we *must* adhere to the way of talking embodied in the theorem; and *this* necessity must be one imposed on us, one that we meet with. It cannot itself express the adoption of a convention; the account leaves no room for any further such convention. (pp.328-9).

The declared aim of the present paper is precisely to explain 'the status of the assertion that certain conventions have certain consequences', and thus to answer both Dummett's questions on the basis of an amplified form of what he calls 'modified conventionalism'. But the wording of his objection suggests, perhaps unintentionally, an argument which might be brought against even a conventionalism which did succeed in reconciling the concept of committal with the behavioural theory of meaning.

The argument runs like this: The amplified form of conventionalism, whatever its details, will have to define 'committal' and then say that such-and-such a sentence expresses a necessary proposition because our committal to it follows from the explained sense of 'committal' together with other relevant facts about linguistic usage; but what about the 'follows from' which occurs in that explanation? The proper answer to this is that conventionalism can explain this 'follows from' too, by a re-application of the conventionalist pattern of explanation to the words occurring in the original explanation. This new explanation will contain a further 'follows from' which can be explained in its turn. There are, it is true, no such explanations that do not contain a 'follows from'; but a cool look at the demand that there should be such an explanation makes it plain that the demand is an absurd one, for it amounts to asking for a derivation of necessity from its source in some way which does not involve the employment of any principles of derivation. I therefore conclude that the aim of the paper is to answer Dummett's, questions or, if it is not, then that the questions admit of no answer.

A current view which I have tried elsewhere to scotch (*Analysis*, Jan. 1961) is that in this century some sort of discovery has been made about the non-informativeness, non-factuality or essential emptiness of logical truths. Now someone might; make the mistake of thinking that he had a triviality doctrine to maintain and that he could maintain it only by adopting a choice-concept of proof and thus avoiding any need to allow that conventions might have remote consequences. There is no reason to think that Dummett makes this mistake . But in Wittgenstein's *Remarks on the Foundations of Mathematics* there is at least one passage in which Wittgenstein shows that he feels pushed towards a choice-concept of proof not by the behavioural theory but by a muddled hunch to the effect that necessary truths must not be allowed to say anything. The crucial remark in this passage is:

> The dangerous, deceptive thing about the idea 'The real numbers cannot be arranged in a series'. . . resides in its making what is a determination, formation, of a concept look like a fact of nature' (p. 56).

I mention this matter here only because the triviality doctrine, like the invalid objection to modified conventionalism discussed above, threatens to become confused with the real issue before us, and must be prevented from doing so by clear identification as a red herring.

3. Conceptual deviation

Few if any philosophers have uninhibitedly adopted a choice-concept of proof. It is obvious to everyone that when a putative proof is presented for the first time and is accepted by a number of competent people, it almost always happens that it is accepted by every competent person who inspects it, including those who do not know that they are thereby mounting a band-wagon; nor is it reasonable to treat this concurrence of opinion as a *sui generis* continuing miracle. Again, it can hardly be denied that when someone is carefully developing a proof of the sort usually called‘ rigorous’, he does not :merely guess or hope that those who accept his premisses will accept his conclusion—he has a strong and in some way justified confidence that they will do so.

Facts like these lead some philosophers, looking for ground between the choice-concept and the committal-concept, to say that what a proof does is to call our attention to the path which, once it has been pointed out, seems to us the most natural one to take. They usually add something to the effect that the conclusion which it is *natural* to draw from a given set of premisses is just the conclusion which one has sound inductive reasons for believing that most speakers of the language would draw. Enquiries as to what these inductively-based reasons are like usually elicit some such reply as ‘I have inductive reasons for believing that if a proof strikes me as valid it will strike most competent people as valid.’ I do not intend to discuss this extravagantly implausible proposal here, as I shall shortly offer a rival positive account of logical committal. But it should be said at once that this gloss on ‘natural’ is wrong at its root, i.e. that whatever it is which leads one to expect certain proofs to be generally accepted as valid, it is not an inductive argument of any sort whatsoever. I shall try to justify this statement in the next section; but first it needs qualification for the following reason.

It is certainly true that when someone is developing a proof, however rigorous, he must admit the possibility that few if any of his audience will regard it as compelling, because of the possibility that most or all of his audience will turn out to be too muddled to follow it. He must also admit the possibility that he has made a mistake in the proof—a straightforward error which he will acknowledge as such as soon as it is pointed out to him—and that for this reason his proof will not be generally found compelling. Now, he can certainly have inductive reasons for believing that neither of these possibilities has been or will be realised; and, to that extent, inductive reasons can be relevant to the belief that one’s own conceptual procedures will continue to be in step with those of others. But such possibilities do not exhaust the range of ways in which it is possible that a proof should fail of general acceptance, and the residue of the ways in which this can happen provide the matter of our problem. It is in respect of the most important sub-division of these, what I call the ‘radical deviations’, that I claim, and shall try to show, that inductive procedures are irrelevant.

Our ‘residue’ contains all possibilities of conceptual deviation which are not attributable to plain muddle or mistake on either side. What is to count as unmuddledness here? I think that the following condition is sufficient, without being much too strong:

> If two or more people continue indefinitely to make the same judgments about what statements are logically true and what proofs are valid, and if they do this without collusion or hesitation and without being involved in clown-like misencounters with the physical world, then they are not confused or muddled. If their judgments about necessity and validity differ

> from ours, we must allow that they are consistently operating a conceptual scheme but that it is not ours.

It is possible that there should be two or more people whose patterns of behaviour force us to admit that they are not in any ordinary sense muddled, that they are in possession of a conceptual scheme (the same conceptual scheme), but that this system of concepts—while coinciding with ours over a large area—differs from ours in unexpected ways. Confronted with such people, we may wish to say 'You cannot give a unitary meaning to (say) the word "and" in accordance with which it is correct for you to use it just as we do on tens of thousands of occasions and also correct to use it in these other, strange, false things that you sometimes say. There isn't any such meaning for the word to have.' But we are not entitled to say any such thing: there is such a meaning, for the pattern of linguistic behaviour defining it is there before our eyes.

I want to concentrate on the special case in which the area of coincidence between our conceptual scheme and the deviators has been such as for a while to mask completely the fact that they are deviators, the case in which they suddenly start using the old familiar words in (what we call) strange new ways while insisting that they are using the same conceptual scheme throughout but that something strange has happened to *our* use of language. For the possibility that this should happen is just the possibility that, while everything seems normal, we are on the brink of finding ourselves at a parting of the conceptual ways. If this is always possible, than how can it ever be correct to say that such-and-such premisses, given all our other uses of the words in them, commit us to such-and-such a conclusion? For to say this is precisely—is it not?—to say that in respect of the matter in hand there is no possibility of deviation.

4. Diagnosing deviations: radical cases

Let us suppose, then, that someone who has taken an apparently normal part in our linguistic community suddenly begins disagreeing radically with us over points of logic and convinces us, by an appeal to one or more others who independently disagree with us in just the same ways as he does, that we have no good grounds for saying that he is trying to employ our concepts but is making a bad job of it. If we decide to look into such a situation exhaustively, what are the sorts of things we can find?

(1) If the deviation concerns descriptive general words—for example, if the deviator denies 'Nothing is blue all over and red all over at once '—then a possible diagnosis is that there are one or more general terms to which the deviator attaches meanings other than the ones we attach to them. He might, for instance, mean by 'blue' what we mean by 'smooth'. This could occur only if everything which he had heard described as 'blue' felt smooth to him, or looked the way smooth-feeling things look; and if all the things he had heard described as 'smooth' looked, say, blue to him. (A precise switch of the meanings of ' blue' and' smooth' is not necessary for this story, merely convenient.) We could discover that by 'blue' he meant what we mean by 'smooth' if (and, nearly enough, only if) it turned out that, although his groupings of things as 'blue' and 'smooth' had so far tallied with ours whenever they were checked against ours, when we really set him to work extensively classifying things under these two labels the vast majority of things he called 'blue' we called 'smooth' and vice versa. There is nothing impossible in this suggestion: the proportion of the world's furniture on which we compare notes is small, and the remainder offers ample opportunity for a breakdown of our community of descriptions, and indeed for a breakdown of such proportions

as to swamp the area of agreement. Adapting the terminology of Quinton's 'Properties and Classes' (*Proceedings of the Aristotelian Society* 1957–58) we could describe this as the belated discovery that some classes which are natural for the deviator are not natural for us, and vice versa. More specifically, this is the very special case in which, say, 90% of the members of a class which is natural for us form a class which is natural for the deviator, the remaining 10% of the members are for him part of a distinct natural class, and it happens that up to a certain time every occasion of checking our use against his use of our word for the whole class has chanced to be an occasion involving some of the 10% and not the 90%. That the deviator should for years not betray his need to determine smoothness visually, and his preference for determining blueness tactually, represents an additional implausibility but not an impossibility.

(2) Suppose the deviator disagrees with us about the logic of words which are not descriptive general terms at all—for instance, if he declares some theorem of the lower predicate calculus to be false on its standard interpretation. It is possible that the explanation here is the same as in the previous case, namely a non-coincidence of the classes which are natural for the deviator with those which are natural for us. This could lead to a divergence in the meanings attached to what we might call 'formal' words, for the meanings of the 1atter are given mainly through their uses in application to the world, and radical deviations in the apprehension of the world might by very bad luck lead to consequent deviations in the understanding of formal words. We might start to develop an example by supposing the non-coincidences between the deviator's natural classes and ours to be such that, until the deviation becomes manifest, almost every time the deviator has been in a position to check a statement of the form 'All S's are P's' it has appeared to him that some S's were P's but not all, and every time he has been in a position to check a statement of the form 'Some S's are P's' it has appeared to him that all S's were P's. This could lead him to attach to 'all' the meaning we give 'to 'some' and vice versa. This switch would involve correlative deviations in his understanding of other formal words, and the postponement of the discovery of these would be possible only by virtue of yet further non-coincidences between the deviator's natural classes and ours. The complexities would have to be, at least, astronomical. But it is not part of my thesis that this is a likely tale, only that it is a possible one. Nor is even that essential to the conclusions I shall reach.

By remainder, we are left with the possibility of a conceptual deviation which our best efforts cannot show to result from any abnormality in basic classifications. Let us suppose that a deviation of this sort occurs involving the sentence-connecting 'and': for years the deviator's use of the form 'P and Q' has obeyed the standard truth-table (I disregard here, as below, the consequences of occasional factual error on the deviator's part); but after the initial deviation and many more it becomes manifest that the deviator's 'and' is not truth-functional at all. Now, we could in principle make as long a list as we liked of the pairs of sentences whose 'conjunction' the deviator is prepared to assert and another of the pairs whose 'conjunction' he is prepared to deny. If the lists are in chronological order of audible assent, then an enormous initial segment (but not the whole) of one list will have the property of conjoint truth while an enormous initial segment (but not the whole) of the other will lack this property. Diagnosis of such a deviation must begin with an attempt to find some property common to all the pairs in one list and none in the other. The different possible outcomes of such an attempt form the bases for the third radical type of deviation and for the range of superficial

types which I shall; discuss' in the next section.

(3) It might be that the best generalization we can find to cover all the pairs in the first list and none in the second is of such horrid complexity that we can hardly believe it to convey what the deviator has in mind, and are accordingly hesitant over using it as a basis for predictions about the deviator's further uses of 'and'. And if things were so bad that we could do no better than a 'generalization' which was only trivially general, being based on a disjunction of identifying descriptions of all the items on one of the two lists, we should be entirely unwilling to use our 'generalization' to make predictions. For we could do so only by choosing between taking the first list as complete and predicting that every new pair of sentence-types would go in the second list, and taking the second list as complete and predicting that every new pair would go in the first. The arbitrariness of either choice, no less than the complexity of the 'generalization' would make it absurd to have any confidence that we had formulated the deviator's concept of 'and'.

The present case must be construed as one in which some principle of conceptua1 organization which is immensely complex for us is simple for the deviator, but not because he is so clever that he can do rapidly and easily what we can do only stumblingly and with difficulty. Suppose, for example, that the simplest generalization we can find to cover his 'conjunctive' list is of the form:

> Any given conjunction occurs in the conjunctive list if and only if both its members are true and if there is a time-reference in the first conjunct and there are two time-references in the second conjunct then the one in the first is temporally further from either in the second than those in the second are from one another) or (the first is true and the second false and. . .) or. . .

and so on, with about ninety parentheses. A capacity for handling this rule as easily as we handle our rule for 'and' cannot be sensibly attributed to abnormal cleverness unless the deviator shows superhuman intelligence in other areas too. In any case it would be a strange sort of high intelligence which succeeded in missing the point of our use of the word 'and'. Facility with numbers will enable a man to see more formulae than we can for continuing the series '1, 3, 5, 7,. . . ', but it will not make it difficult for him to think of the continuation '. . . 9, 11, 13,. . . '. It seems fairly clear that the hypothesis of superhuman capacity can have no place here except perhaps as an irrelevant addition.

If a public description of the world is to be possible for a group of people, the members of the group must roughly agree in what classes they find natural, but it is also necessary that they should roughly agree in what sorts of conceptual organization they find relatively simple and what sorts. they find relatively complex. This amounts to saying that the rarity of deviations of type **(3)**, just like the rarity of deviations of types **(1)** and **(2)**; is a pre-condition of the possibility of applying a public language to the world in even a primitive way. The point about **(1)** and **(2)** is obvious—or if it is not, see Quinton, *op. cit.*—but the point about **(3)** is perhaps less so. Consider someone trying to pick up the meaning of the word 'purple' from its application to particular purple things: he may of course go wrong because purple things do not form a natural class for him, but he may go wrong in another way too. It will always be the case that all the. things called 'purple' and none of those denied to be 'purple' will answer to some long disjunctive description each disjunct in which is a conjunction of perfectly ordinary descriptions (like 'purple'). Now, if the person trying to learn the meaning of 'purple' regards the fact that a pair of things answer to that description as being at least as likely and useful a basis for saying they 'have something in common'

as would be provided by their being purple, then he may continue to use 'purple' on the basis of whether they answer to that description rather than on the basis of whether they are purple. That this kind of thing does not occur, although often enough the required description could be formulated in quite ordinary and manageable terms, is due to the fact that we should find it difficult to think of such a description in the first place, still more difficult to keep it in mind for any length of time, and impossibly difficult to re-apply it to the world confidently and rapidly. To explore in the detail it deserves this possible source of failure for any ostensive procedure would, 1 conjecture; be to discover that there is an element of fiction in the very assumption that there is any clear distinction between non-coincidence of natural classes and non-coincidence of judgments about relative conceptual simplicity. For present purposes, however, what matters is that even if we falsify the situation to the extent of assuming the two to be easily distinguishable, and thus separating **(3)** from the first two types, we still get the result that **(3)** relates in just the same way as **(1)** and **(2)** do to the possibility of any communication.

We are now in a position to see why inductive procedures are irrelevant here. If I had doubts as to whether a conclusion I was inclined to draw would be drawn by all intelligent speakers of English, I should consult a better than random sample of them, and in time I should have all the evidence of such an inductive kind that I could possibly want; but I should be no nearer to being satisfied as to the unlikelihood of my meeting with a radical deviation. That possibility is just the possibility that inductive procedures should become irrelevant to determining a proof's chances of acceptance.

This does not mean that we have better-than-inductive reasons for believing that radical deviations will not occur; any more than we have better-than-inductive reasons for believing that inductive reasons (of any sort, on any subject-matter) will continue on the whole to lead to true conclusions. All we have is the assurance that the possibility that radical deviations should occur, like the possibility that the world should cease to be amenable to explanation in accordance with laws, is a *dead* possibility: we can do nothing to guard against its realization; we could do nothing to remedy any situation arising from its realization; and it is not possible even to describe or conceive of its extensive realization—for only if it is not realized extensively are descriptions possible at all.

As a footnote to this section, mention should be made of Hampshire's *Thought and Action*, p. 31:

> Everything resembles everything in some respect. . . We could go on picking out resemblances for ever; inexhaustibly. . . Reality by itself sets no limit. The limit is set by changing practical needs and by the development of new powers and new forms of social life.

I think it is fair to say that Hampshire's arguments for this view in fact support only the statement that reality does not set all the limits, and do not come near showing that it sets no limits, i.e. showing that any agglomeration of objects could be regarded by humans as the sole bearers of some property. This stronger statement seems to me to be manifestly false; I have tried in this section to show that it is necessarily false.

5. Diagnosing deviations: superficial cases

Superficial deviations are ones in which—to continue with the example of 'and' —we look over the deviator's conjunctive list and find that all the pairs of sentences on it, and none on the other list, are characterized by some property which is so straightforward and elementary that we have no hesitation in

concluding that we have found the deviator's concept of 'and', or in predicting his further uses of the word accordingly. (Any such prediction may go wrong: the deviator's conjunctive list may admit of simple generalization at noon, only to become chaotic by lunch-time; on the other hand, it may admit of simple generalization at noon and be destined to go on answering to that generalization throughout every addition which the deviator makes to it. All we can meaningfully consider is what we find at the time at which we find it.)

Regarding the distinction between radical and superficial deviations: the main thing is that radical deviations do, while superficial deviations do not, involve the existence of people whose basic abilities and disabilities are out of line with those of humans in general. It must be granted that there is a sliding scale with superficial cases at one end and type-**(3)** radical cases at the other; which is only to say that complexity, or difficulty, is a matter of degree. The existence of intermediate cases does not affect what I wish to say about the two ends; we have only to remember that there are many possible cases which would have elements of both the radical and the superficial. Such a mixture is comparable with the mixture we are confronted with when we think about the possibility of a situation which hovers between chaos and obedience only to extremely complex laws.

Before coming to a superficial deviation involving 'and', let us consider the following case. Someone says 'There are horses with no heads', and it turns out that he calls 'horse' any object which could transport a human sitting astride it. The rule for his use of 'horse' is thus perfectly manageable: it is even shorter than our rule for the use of 'horse', and relates to ours roughly as genus to species. But even if it were not so related—if for example he applied the word 'horse' to all horses and all grapefruit and nothing else—we should still be able to find a fairly manageable rule for his use of 'horse' by disjoining our own rule with a rule for his extension of the sense of 'horse', provided that the latter rule was itself manageable.

Any descriptive general word is vulnerable to such treatment at the hands of someone who has used it, and has been heard to use it, only in particular affirmative applications. A word's openness to this kind of deviation cannot be reduced merely by the frequency and scope of such applications: a deviator who had heard every horse in the universe described as a 'horse' could still go on to classify motor-cycles or grapefruit as horses without prejudice to the superficial status of his deviation. The only, kind of linguistic behaviour which can reduce the likelihood of anyone's giving a word a wider meaning than most people give it, without this first appearing, is the negative application of the word to parts of the world. For example, the extension of 'horse' to mean 'object which could transport a human sitting astride it' is ruled out by the acceptance of just one remark to the effect that a certain motor-cycle is *not* a horse. A special case of this is the giving of necessary conditions for the affirmative application of a word, where words occurring in the conditions have themselves been publicly used in a wide range of negative applications. The deviation sketched above is not ruled out by the deviator's assent to 'All horses are animals' if all he knows of the use of 'animal' is that certain things are animals; but if he learns, say, that motor-cycles are human artifacts and that humans never make animals, then it now becomes impossible for him to extend the sense of 'horse' in that particular direction. And if he accepts a rule saying that all horses are four-legged, then the deviation is almost bound to be ruled out for him, because the language is too full of public negatives involving numbers and legs for anyone to escape enough of them to be able, for instance, to describe two wheels as four legs on the basis of sheer

addition to the uses with which he is familiar.

I have been describing what might be called 'deviation by generalising the normal sense' ('horse' meaning 'object which could transport. . . ' etc.) and 'deviation by disjunction of something with the normal sense' ('horse' meaning 'horse or grapefruit'). There remains only 'deviation by substitution of something for the normal sense', where the deviator is enabled to give an eccentric meaning to 'horse' because everything he has known to be called 'horse' has just *happened* to have some other irrelevant property as well, a property which is noticeable enough in itself but which occurs very rarely or *happens* to have occurred rarely in the deviator's experience. For example, every horse which the deviator is told is a horse happens to be lame at the time, or to look lame to the deviator, and he sees no non-equine instances of lameness until the memorable day when he expresses sympathy with his aunt over her having fallen down and become a horse. Unlike the first two sorts of superficial deviation, this one's chances of occurrence can be reduced by the scope and number of purely affirmative uses of the word in question; in the example, the deviation would have been completely ruled out by a single application of 'horse' to a non-lame horse, just as well as by the aunt's application of 'non-horse' to herself.

Of the three sorts of superficial deviation, those that consist in generalising and substitution can be rendered overwhelmingly unlikely to occur in respect of any given word; it is, for example, unthinkable that anyone should. belong for many years to our linguistic community without having lost every possible chance of varying the meaning of 'and' in either of these two ways. Deviation by disjunction cannot be squeezed out in the same way, as is illustrated by the ease with which we can describe a deviation by disjunction involving the word 'and'. Suppose that a hitherto normal conjoiner deviates by conjoining not only pairs of true sentences but also pairs of sentences of which one is false and the other is true and contains the words 'the Eiffel Tower'. There is no great implausibility in the suggestion that a normal literate adult might escape ever hearing the outright denial of a conjunction one of whose conjuncts is a true sentence containing the words 'the Eiffel Tower'. Even if we think of the deviator as having been almost sure to have assented to something like 'It is incorrect to put "and" between two sentences and assert the result unless the are both true', there may well be nothing with which he is familiar in the uses of words which positively clashes with the interpretation of '. . . they are both true' to mean what we mean by '. . . they are both true or one of them is true and mentions the Eiffel Tower'. In general, there will always be some words with which a given word has not been explicitly linked in the relevant ways, words whose infrequency of occurrence makes them available as irregularities in the smooth logic of other words. So deviation by disjunction is always possible—a fact to which I shall return.

First, there is an aspect of superficial deviations as such which should be discussed and dismissed. It is that if some individual or small group of individuals indulges in a superficial deviation, he or they will have made a plain *mistake*. The superficial deviator errs, as the radical deviator does not, in that we can not only vote him down but we can also show him what his mistake is and can thus bring him to admit that it is a mistake and undertake not to repeat it. What makes superficial deviations theoretically interesting is the possibility that one should occur in circumstances which made 'mistake' inapplicable simply on statistical grounds: a corrigible divergence is only a *mistake* if there is a clear majority to say whose mistake it is.

Any such situation would be rectifiable, of course, but that is just the trouble. The possibilities of radical deviation are dead possibilities not only because we cannot guard against them but also because nothing could be done about them if they were realised: they relate to the question 'Do these premisses force this conclusion?' in about the same way as the possibility of chaos from tomorrow onwards relates to the question 'Does this experiment confirm this hypothesis?' We make no room within the communication-game for the possibility that the game will become unplayable, just as we make no room within science for the possibility that science will cease to be a possible kind of activity. But the possibilities of superficial deviation are not like this at all.

Nevertheless, we can do more than beat our breasts. Consider first deviation by disjunction : this is a 'live' possibility, and it is also a real possibility in the sense that it is always thoroughly on the cards in any society in which words are constantly being learned. But if a deviation occurred which really was of this sort, and was not a veiled case of naturalness-deviation, then the deviator would have to grant that in his sense of it the word concerned was ambiguous to a greater degree than in our sense of it. We could show him that his use of the word required the learning of rules for our use of it plus, as a sheer addition, the learning of rules for his use of it; we could show him that his use of the word makes it, in that sense, tantamount to two or more words spelled in the same way. (This is what marks off 'disjunction of something with the normal sense' from ' from 'true generalization of' the normal sense'.) The point of bringing this home to him would not be to denigrate ambiguity as such but to show him that, by giving a word more ambiguity than the linguistic pressures upon him had forced him to give it, he had assumed the existence of an element of complexity for which he had no evidence at all. We could sensibly say to him, 'Do not in future do that sort of thing.'

Nothing of that sort can be said about deviations by substitution and by generalization. Both of these arise from the intersection of normal learning techniques with plain bad luck, and they cannot be guarded against by a general rule. On the other hand, the likelihood that there should actually occur a deviation of either of these sorts can be reduced at will for any given word. Does anybody think that for many of us it is now a possibility at all in respect of, say, the word 'and'? It does make sense to speak of completely eliminating someone's chances of indulging in a deviation of either of these sorts; for it is not true that the possibilities are endless, even if we can never be sure that we have come to the end of them. It is just possible that John Doe's linguistic history has been such as to leave him room for a superficial deviation, by generalizing or substitution, in respect of the word 'and'. But to say this is not like saying 'It is just possible that there are extra-terrestrial parliamentary democracies', but is like saying 'It is just possible that for the last ten years I have been shadowed by an extremely clever detective'. In the latter case, but not the former, there is a sense for conclusively settling the matter in the negative.

6. Conclusion

I have been primarily concerned to get the question of conceptual deviation into an area in which we can lay it out and inspect it in detail; for this reason I regard the attempt to classify deviations as one which it is important to make. Despite a variety of doubts about the results of the inquiry and about the classification on which it has been based, I am inclined to give credence to the following conclusions:

(1) We can distinguish in a non-question-begging way

between the deviations which can be corrected by the means of normal communication and those which cannot. The latter are those which, not entirely expectedly, turn out to have important points in common with possible kinds of failure of ostensive procedures. If we are to have a proper understanding of this matter, we must beware of such false analogies as the one Wittgenstein draws between the possibility of a radical deviation and the possibility that a machine should not behave according to plan because of 'distortion of parts'; the proper comparison is with the possibility that a machine should just misbehave, for no reason at all.

(2) Given the distinction between radical and superficial deviations, a sense can be found for 'commit' in which the acceptance of premisses can commit us to (inconsistency or) the acceptance of a conclusion, in at least as strong a sense as that in which stepping off a cliff commits us to falling down it. In each case, we can admit the possibility that there could be someone who accepted the premisses and denied the conclusion without inconsistency (who stepped off the cliff and, miraculously, did not fall), but in neither case can we do anything about such a possibility either before or after its occurrence—such a possibility cannot be in any way allowed for in our logic (physics). From this it follows, *a fortiori*, that it is hopelessly wrong to talk of rigorous proofs in terms of decisions, or as procedures which guide but do not force us. The traditional picture of proof which has been associated with the platonic-proposition account of meaning is an entirely correct picture; it is simply a mistake to say that if the behavioural theory is accepted then the outlines in the picture must be in some way softened and blurred.

(3) Let us call any argument 'absolutely rigorous' if it is impossible for any speaker of the language in which it is expressed to accept its premisses and decline to accept its conclusion without either (a) being in a muddle, (b) being guilty of what he will himself call a 'mistake' once it has been explained to him; or (c) being incapable of communication with the rest of the human community at least over a limited area. Some arguments are absolutely rigorous, and some are not. An argument may be wrongly thought to be absolutely rigorous, even by someone who is neither foolish nor ignorant. And someone may wish to hedge his bets by saying that he is not completely satisfied of the absolute rigour of an argument: but we can afford to leave him to his uncertainty, to his bit of Descartes's demon, just so long as he distinguishes 'I am not completely satisfied of the absolute rigour of any particular argument' from 'It may be that no actual argument is absolutely rigorous', and distinguishes both of these from 'The behavioural theory rules out the very concept of an absolutely rigorous argument'. Of these three statements, the first is unduly sceptical, the second is neurotic, and the third is a philosophical mistake.