# 6

# REASONING AND ITS USES

LET me quote, in place of an introduction, what Tolstoy says in *War and Peace* about Platón Karatáev, the old peasant: in this single paragraph is contained the germ of all that we shall discover.

> Sometimes Pierre, struck by the meaning of his words, would ask him to repeat them, but Platón could never recall what he had said a moment before, just as he never could repeat to Pierre the words of his favourite song: 'native' and 'birch-tree' and 'my heart is sick' occurred in it, but when spoken and not sung no meaning could be got out of it. He did not, and could not understand the meaning of words apart from their context. Every word and action of his was a manifestation of an activity unknown to him, which was his life.... His words and actions flowed from him as evenly, inevitably, and spontaneously, as fragrance exhales from a flower. He could not understand the value or significance of any word or deed taken separately.[1]

## 6.1 *Widening the Problem: What is 'Reasoning'?*

'What kinds of argument, what kinds of reasoning is it proper for us to accept in support of moral decisions?' In all our attempts to answer this question up to the present, we have begun by asking, 'What is goodness?' or 'What is rightness?' This approach, for reasons which we have examined, has got us nowhere. Let us now try approaching the question from the other end—asking instead, 'What is "reasoning"?' If we can come to understand reasoning in general, we may be better prepared to solve the special problems of ethical reasoning.

What is 'reasoning'? Let us take a look at a number of typical cases of reasoning. The circumstances in which we talk of 'reasoning', and of 'reasons' being offered in support of conclusions, are so many and so diverse that it is hard to see what is common to them, and the choice is embarrassing. I shall therefore begin by giving four examples chosen from very different

[1] Tolstoy, *War and Peace*, bk. XII, ch. 13; tr. Maude, World's Classics Edition.

types of reasoning, but each characteristic of its kind. I shall present them in the form of dialogues:

(i) *An arithmetical example*

*A.* 'You should have at least four shillings left.'

*B.* 'Should I? I thought I'd spent more than that.'

*A.* 'No. You began with fifteen shillings. Then we took a bus-ride, which cost us a shilling; then we had tea—that was three shillings; you bought a five-shilling book, and finally spent one-and-ninepence at the grocer's. One, and three, and five, and one-and-a-bit makes ten-and-a-bit: eleven from fifteen leaves four....'

*B.* 'Oh yes, so I should. Well, I seem to have dropped half-a-crown somewhere.'

(ii) *An example from science*

*A.* 'Why, of course the fuse would blow if you tried to use that bulb.'

*B.* 'I don't see why. After all, it was a 5-amp. fuse.'

*A.* 'Yes; but the bulb was only a 50-volt one.'

*B.* 'I know, but it was marked 100 watts; and a 100-watt bulb should only take half-an-amp. off 200-volt mains.'

*A.* 'Ah, but the marking only applies if you're using it with 50-volt mains. If you put a 50-volt bulb across the 200-volt mains, the filament immediately overheats and melts, and the surge of current passing through it is quite enough to blow a 5-amp. fuse.'

*B.* 'Yes, I suppose it would be. But let's get on and find a torch; then we'll be able to see where we are.'

(iii) *An example from ethics*

*A.* 'Jones is fundamentally a good man.'

*B.* 'Why do you say that?'

*A.* 'His harsh manner is only a pose. Underneath, he has the kindest of hearts.'

*B.* 'That would be interesting, if true. But does he ever manifest this kind heart of his in actions?'

*A.* 'He does. His old servant told me that Jones never uttered an unkind word to her, and recently provided her with a luxurious pension. And there are many such instances. I was actually present when (etc.)....'



*B.* 'Well, I confess I do not know him intimately. Perhaps he is a good man.'[1]

(iv) *An everyday example*

*A.* 'Come and have a drink!'

*B.* 'We can't go yet. What's the hurry, anyway?'

*A.* 'Don't you know? It's my birthday.'

*B.* 'Many happy returns, old man; but what difference does that make?'

*A.* 'The boss said I could pack up early and go and celebrate. I'm sure he wouldn't grudge me your company.'

*B.* 'Fine. In that case, give me a moment to put my things away and I'll be with you.'

The most obvious thing that these examples have in common is the form of their dialectic. In each case, one of the speakers, *A*, starts off by saying one thing ($a_0$). In each, the other speaker, *B*, wavers, is sceptical or positively disagrees (saying $b_0$). *A* goes on to make a different remark ($a_1$). *B* is, perhaps, still unconvinced, and does not yet agree ($b_1$). The conversation continues, *A* presenting a series of fresh considerations ($a_2, a_3, \ldots a_n$) and *B* still wavering or disagreeing ($b_2, b_3, \ldots b_{n-1}$); until, at the end, *B* agrees ($b_n$) not only to *A*'s last remark ($a_n$) but also to his first one ($a_0$) and, in many cases, to all the intermediate ones ($a_1, a_2, a_3, \ldots$) as well.

Even if the two speakers were using a language so strange that we could do little more than recognise whether two remarks were the same or different, we might very well surmise, if their conversation had this dialectical form, that *A* was 'giving reasons' to *B*, and that the utterances $a_1$ to $a_n$ were the 'reasons' for his original utterance $a_0$.

If this is what all the cases of 'reasoning' have in common, can we perhaps *define* 'reasoning' as argument having this dialectical form, and 'reasons' as those utterances occupying the places $a_1$ to $a_n$ in a 'reasoned argument' (so defined)?

[1] Cf. Stevenson, *Ethics and Language*, p. 29.

## 6.2 *'Gerundive' Concepts*

The answer is that we cannot: the dialectical pattern is too wide. Although the most typical dialogues in which 'reasons' are offered for a conclusion do fit it, so do dialogues of other kinds, ones which are emphatically not instances of 'reasoning'.

If we are to conclude that $a_1, \ldots a_n$ are 'reasons' for $a_0$, it is not enough that *B* should end up by saying what *A* wants him to. He might very well do the same in response to threats, or taunts, or jeers:

*A*. 'You call me that again!'
*B*. 'Shan't.'
*A*. 'Go on! I don't believe you dare.'
*B*. 'Shan't.'
*A*. 'Oh, won't you? Cowardy-cowardy-custard!...'
*B*. 'All right; I shall then—You're a beastly bully!'

Again, for $a_1, \ldots a_n$ to be 'reasons' for $a_0$, it is not enough that *B* ends up by agreeing sincerely and genuinely to what *A* wants. There are all sorts of situations in which a conversation of the right form may lead to this result, and $a_1, \ldots a_n$ may still not be things we should call 'reasons':

*A*. 'Come and have a drink!'
*B*. 'I can't: it's not time to go yet.'
*A*. 'Oh, come along! Don't bother about that.'
*B*. 'O.K. I don't mind if I do.'

If it comes to that, it is not enough that *B* should end up by saying, accepting and *believing* what *A* wants him to. His judgement over the question at issue may be faulty. He may be misled into accepting a mathematical conclusion by taking a special case as proving a general theorem. Or, if $a_0$ is a scientific hypothesis, he may be too readily impressed by experimental evidence which is in fact inadequate or irrelevant. If the argument is an ethical one, he may take 'everybody does it' as a reason for adopting a pernicious habit. And there is a multitude of specious excuses for having a drink, by which he may be led astray at any time of day or night.

As long as we confine our attention to what *A* says, or *B* agrees, or *C* believes, we shall fail to find the answer we want. And this shows the weakness of our approach: it is not the form of the dialectical pattern, whether alone or in conjunction with the attitude of the speaker and hearer, that makes the utterances $a_1, \ldots a_n$ 'reasons'.



When we talk about the 'truth' of a proposition, or the 'validity' of an argument, we are interested in something which applies to the proposition or argument regardless of who believes it. To conclude that a proposition is true, it is not enough to know that this man or that finds it 'credible': the proposition itself must be *worthy of* credence. Likewise, in order to decide that an argument is valid, we cannot rely on the fact that such-a-one regards it as 'plausible': the argument itself must be *worthy of* acceptance, as making its conclusion *worthy of* belief. And the same kind of considerations apply to all logical concepts—'correct', 'sound', 'relevant' and so on.

Questions of logic, that is to say, are concerned, not with 'subjective relations'—with what is 'credible' (to-*A*), 'plausible' (to-*B*), etc.—but with concepts of a *different* kind.

In this, they resemble questions of ethics (and, for that matter, of aesthetics). I have analysed at length the 'subjective' approach to ethics, which attempts to account for ethical concepts in terms of the attitudes of the speaker and hearer alone, and shown that it is fallacious. If we are to conclude that some past deed was 'good' or that some proposed course of action is 'right', it is not enough for us to know that we ourselves are psychologically disposed to approve of the deed, or that the proposed course seems right to the agent: we must have reasons for thinking that the act was *worthy of* approval, or that the course of action is *worthy of* selection. (And, in the same way, for a picture to be 'beautiful', it is not enough for it to attract *me*: it must be *worthy of* admiration.)

Questions of ethics and aesthetics, as well as of logic, are evidently concerned, not with 'subjective relations'—with what is 'attractive' (to-*A*), or with what 'seems right' (to-*B*)—but with concepts of a *different* kind.

These concepts—logical, ethical and aesthetic alike—we can class together as 'gerundives', thereby opposing them to such logical categories as 'properties' and 'subjective relations'. The name 'gerundives' is appropriate because they can all be analysed as 'worthy of something-or-other'; in this resembling

the grammatical class of 'gerundives', which appears in one's Latin primer—consisting of such words as *amandus*, which means 'worthy of love' (or 'meet-to-be-loved'), and *laudandus*, which means 'worthy of praise'.

In the light of this distinction (between 'subjective relations' and 'gerundives', between the 'credible' and the 'true') we can make more explicit the problem before us—the problem we expressed vaguely at the beginning of this chapter in the question, 'What is "reasoning"?' What we have to discover is why some of the arguments which fit our dialectical pattern deserve the title of 'reasoning' and some do not. We must therefore ask: 'Of what kind do the utterances $a_1$ to $a_n$ have to be in order to make the conclusion, $a_0$, worthy of acceptance, in a way in which others do not?'

Put in this way, the present problem clearly resembles our central problem. It differs from it, in fact, only in being more general—in applying to arguments of any kind, and not simply to those from factual reasons to ethical conclusions.

### 6.3 *Philosophical Theories of Truth*

One might have supposed, from a quick look at this question, that the answer depended entirely upon the kind of conclusion to be justified, and upon the circumstances of the case. One might have supposed, that is, that the kinds of thing which make a scientific conclusion worthy of belief were so different from those which make (say) a mathematical, ethical, psycho-analytic or aesthetic conclusion true, that nothing valuable could be said which would hold for them all. And one might have doubted, in consequence, whether there was any hope of finding an explicit answer to the question, since any verbal formula comprehensive enough to cover all modes of reasoning must be so vague as to be useless.

Once again, however, the philosophers claim to know better. For most of them, our fears and doubts are unreal; and they are ready to produce, in answer to our question, a comprehensive verbal formula of the kind which seems to be required. To tell the truth, they produce between them a large number of such formulae, many of them at first sight appearing to be incompatible with one another. In considering the question we are, in



fact, in the position of having too many answers to choose between rather than too few.

Some of the philosophers declare that, to make the conclusion worthy of belief, the arguments must show 'that it corresponds to a fact'.[1] Others declare, instead, that they must show 'that it coheres with our other beliefs'. A third school—'that it is of benefit to those who use it'. A fourth—'that the members of the proletariat believe it'. And these are only a few of the answers given.

These 'philosophical theories of truth' are comparable with the 'philosophical theories of ethics' discussed in our earlier chapters. They may be divided into two classes, according as they regard the 'truth' of the conclusion as a property of the conclusion itself—these we may naturally call 'objective' doctrines—or as concerned with the attitude, interest or welfare of the speaker, or some limited group of people to which he belongs —that is, 'subjective' doctrines.[2] (In some ways, the thing which surprises me most is that I have never come across an 'imperative doctrine of truth'; especially considering that it would be easy and entertaining to develop one—and quite a plausible one, too. Could one not argue, with considerable force, that to say that a conclusion is 'true' is not to attribute a 'property' to the conclusion, or to express one's 'attitude' towards it, but is just to evince one's belief in it, and to bring pressure to bear on one's hearers in the hope that they will assent to it and believe it as well?)

The question we must ask ourselves, at this point, is whether such philosophical theories of truth—and, for that matter, the corresponding theories of aesthetics, the 'objective', 'subjective' and 'imperative doctrines of beauty'—can hope to succeed in distilling the essences of 'truth' and 'beauty' into comprehensive verbal formulae.

If we hark back to our earlier discussion of the theories of 'goodness', and recall the reasons why the philosophical theories of ethics failed to do, for 'goodness', what these theories want

[1] See John Wisdom, *Problems of Mind and Matter*, ch. XI.

[2] The distracting effect of spatial metaphors can be shown as well in the 'theory of truth' as in philosophical ethics (cf. §§ 3.8 and 4.5) and metaphysics (cf. § 8.9). Recall the old tag, *Judicium est 'locus' veritatis.*

to do for 'truth' and 'beauty'; and if we remember also how it was that we came to class 'beautiful', 'true' and 'good' in one category as 'gerundives'; then our answer to this question must be 'No!' The arguments which led us to reject all three types of 'ethical theory' could be applied equally well to the 'theories' of truth and beauty. And they would demonstrate, as rigorously as they did for the theories of philosophical ethics, that such philosophical doctrines of truth and beauty cannot succeed.

In consequence, we may conclude:

(i) That the questions, 'What is truth?' and 'What is beauty?', if answered directly, are no more fruitful than the corresponding questions, 'What is goodness?' and 'What is rightness?'

(ii) That all the short answers given to these questions are, if taken literally, false; and that, if taken figuratively, they can at the best only focus attention on some special feature of the concept, all-important, perhaps, over a limited range of instances, but not of universal application.

(iii) That the central practical questions, 'What kinds of thing make a conclusion worthy of belief?' and 'What kinds of thing make a work of art worthy of admiration?', are to be answered, not by verbal pantechnicons with room for every case, but by a discussion of the ways in which, when faced with some particular variety of sentences, or works of art, we should set about making our selection.

It would be tedious to repeat in detail for 'true' and 'beautiful' the arguments we have already given for 'good'. We can be content with illustrating the force of our conclusions as they apply to a typical 'theory of truth'. I shall choose, for this purpose, a theory which is at once the most fashionable, and in some ways the most plausible—the 'correspondence' theory of truth.

## 6.4 *The 'Correspondence' Theory of Truth*

According to the 'correspondence' theory of truth, to say that a proposition is 'true' is to say that it 'corresponds to a fact'. (This is to locate the 'truth' of the proposition 'in' the proposition itself, rather than 'in' the speaker or hearer, so the theory may be classed as an 'objective' one.) Supporters of this theory argue that the utterances $a_1$ to $a_n$ can only make the conclusion, $a_0$, 'true' if they show that it 'corresponds to a fact'.

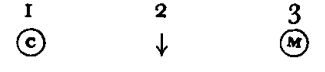


Now, to do justice to the theory, and to its advocates, it does seem to give a life-like picture of what we require of certain types of utterance when passing them as 'true' or 'false'. There are some sentences which we can describe, with almost literal vividness, as 'corresponding to', or even as 'giving a picture of', those features of the world which they describe. Sentences of this kind (we may say) have a 'structure': that is, the sentence can be split up into a number of elements; each of these elements 'refers to' something in the world; and the mutual relations of these elements, in the sentence, are like (if the sentence expresses a 'true' proposition) or unlike (if it expresses a 'false' one) the mutual relations, in the world, of the things to which the elements 'refer'.

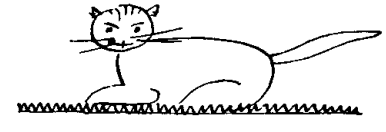
Consider, for example, the sentence 'The cat is on the mat', which appears in a child's first reader. This sentence can be split up into three elements thus:

'1 (The cat) 2 (is on) 3 (the mat)'

or, more symbolically, thus:




The situation which this sentence describes may be drawn thus:



and this state-of-affairs, fact, or what-you-will can be represented, in its turn, thus:


When one allows for the conventions governing the writing of the English language[1]—that it shall consist of horizontal rows of

[1] One can imagine a language in which these conventions were less rigid—in which, for example, spatial relations were expressed pictorially, so that one wrote

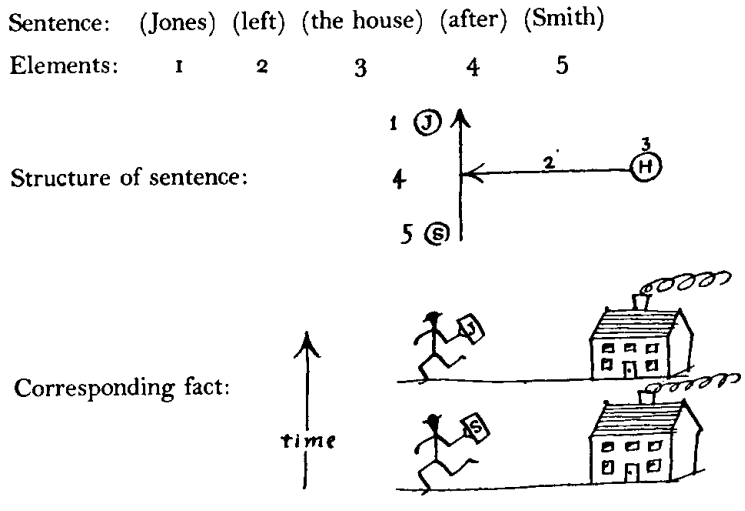
'Cat
mat'

instead of 'The cat is on the mat'.

One can even imagine a people among whom photographs or drawings took the

symbols, and so on—it will be clear that we can regard the 'structure' of the sentence 'The cat is on the mat' as reflected in the 'structure' of the situation which it describes, but as different from that of any of the situations falsifying the statement.

We can do the same kind of thing with other simple descriptive sentences. Suppose, for example, that I say 'Jones left the house after Smith'. Then, if my statement is true, the temporal order of the words 'Jones' and 'Smith' is directly related to the order in which the two people referred to passed through a certain doorway. And, if it is false, this correspondence between the elements of my sentence and of the situation described is absent. Again, we might display the things which make the sentence true diagrammatically, thus:




And again the 'truth' or 'falsity' of the sentence depends on there being a 'correspondence' between two 'structures'—the 'structure' of the elements of the sentence and the 'structure' of the things in the world to which the elements 'refer'.

(Incidentally, this second example underlines an important point which is obscured in the case of the first. In looking for the 'correspondence' between the two 'structures' in any particular

place of descriptive sentences—so that the drawing I have included in the text was used as a 'hieroglyph' or 'pictogram', taking the place of the sentence. But on this point, see §6.5 *ad fin.*



case, we require a rule of interpretation specifying the exact nature of the 'correspondence'. We have, for example, to learn that, when the spoken sentence 'Jones left the house before Smith' is true, the events referred to have the same temporal order as the words 'Jones' and 'Smith', whereas in the case of 'Jones left the house after Smith' the temporal order of the events is reversed.)

The 'correspondence' theory of truth does, therefore, give a vivid and illuminating account of one feature of the truth of a certain range of sentences, a range which could be made much wider than I have done by examining in more detail the ways in which words are made to 'refer to' things. In this, the theory resembles the philosophical theories of ethics: but, in order to bring out the full force of the analogy between 'truth' and 'goodness', I must also show that it resembles them in another respect —that, if the 'correspondence' theory is assumed to be of universal application, its consequences are paradoxical and even nonsensical.

### 6.5 *Correspondence and 'Description'*

It is natural enough, if one is looking for a single, exhaustive answer to the question, 'What is truth?', to choose as that answer 'correspondence with fact'. The vividness of this phrase, in the cases to which it most exactly applies, commends it above other possible answers; and it is not surprising to find that those who have set out to 'exhaust the universe of logic' (for example, Wittgenstein, in his *Tractatus Logico-Philosophicus*) have taken it as their starting-point.

The failure of such attempts is, however, guaranteed before they begin. It is implicitly assumed in any statement of the 'correspondence' theory of truth that all our significant utterances are intended as *descriptions*. I could not have explained the theory without giving a 'descriptive sentence' as an example, and without talking about the features of the world to which it 'corresponds' as those which it 'describes'. Indeed, the fact of the matter is that, if one starts off from the 'correspondence' theory of truth, the only logical category one can 'exhaust' is that of 'descriptive sentences': all other types of utterance slip through the net.

When Wittgenstein wrote, in the *Tractatus*: 'The specification of all true elementary propositions describes the world completely. The world is completely described by the specification of all elementary propositions plus the specification, which of them are true and which false'[1] (where an 'elementary proposition' is one which 'describes' and 'asserts the existence of' an 'atomic fact'), his explicit use of 'describes' and 'described' should have made this limitation doubly clear. And yet, at the same time, he seems to have supposed that his remarks were relevant to logic as a whole, for he went on to remark that many of our utterances cannot be regarded as expressions, either of elementary propositions, or of combinations of them, and concluded that such utterances are not logically respectable. 'There are', for instance, according to the *Tractatus*, 'no ethical propositions.'[2]

All that one can prove about ethical utterances by these means, however, is that they are not descriptive: and this is not news. The first step in applying the 'correspondence' test for truth is to split the sentence concerned into elements, each of which 'refers to' something in the world:[3] until one has done this, all talk about 'comparing structures' or 'looking for a correspondence' is useless. But, when we are interested in ethical sentences, the instruction to 'split' the sentence into 'elements' of this kind can mean nothing to us; for ethical concepts, as we saw when we discussed the objective and subjective doctrines of 'goodness',[4] cannot be said to 'refer to' anything at all, either in the speaker or in the world about him. In consequence, the 'correspondence' theory, so vivid and life-like as applied to one range of utterances, becomes nonsensical and obstructive when applied to another, equally 'respectable', but inappropriate range. And if the criterion of truth that it emphasises does not make sense when applied to at least one range of utterances, what can be the use of exalting it into a universally applicable theory of truth?

[1] Wittgenstein, *Tractatus Logico-Philosophicus*, 4.26, p. 91.

[2] Op. cit. 6.42, p. 183. I do not want to suggest that Wittgenstein still holds the views quoted here: I am quite certain that he would be prepared to criticise them at least as severely as I have. Still these quotations have a great *historical* importance.

[3] As we did in § 6.4 with the sentence: (The cat) (is on) (the mat).

[4] See § 3.8 above.



Is there any way of explaining why the 'correspondence' criterion of truth should apply only to descriptions? I believe there is. For consider what we do when we make up a description. Suppose, for example, that 'Slimy Joe' has been given ten years' hard for forgery, and that he arrives at the jail. One of the first things now to be done is that 'full particulars' will be taken of him: that is to say, he will be stood in the light and a warder will make a note of the colour of his hair—'fair'—and his complexion —'ruddy'—his height, chest circumference and other dimensions will be measured—'6 ft. ½ in., 35 in. expanding to 39 in.', etc.— and any special features will be noted down—'scar over left eye'.

If, now, 'Slimy Joe' ever comes to escape, or goes back to his old occupation after his release, a 'Wanted' notice will be prepared for distribution to all police stations. The purpose of this notice will be to bring about his *recognition*, and so his apprehension. It will, therefore, give a full 'description' of him, based on the prison records—'6 ft. ½ in. in height, fair hair, ruddy complexion, with a scar over the left eye', etc. In so far as this is a good description, it will help its readers to recognise him: if, however, by some mistake, someone else's particulars get sent out instead—'About 5 ft. 4 in. in height, sandy hair, pasty complexion'—recipients can justifiably complain that no one could be expected to recognise 'Slimy Joe' from *that* description. In so far as the circulated notice helps people to recognise the wanted man, it is a good description: in so far as it fails to assist this recognition, it fails *as a description*.

This gives us our clue. In composing a description, we have to produce an utterance *corresponding recognisably to* whatever it describes: in verifying a description, we have to confirm that it does *correspond to* the thing described. The reason why the 'correspondence' criterion of truth applies so aptly to descriptive sentences is, therefore, because with its help we can discover whether they have *served their purpose*—and, if the rules that it gives for verifying descriptive sentences look like the rules for giving a description 'in reverse', is that at all surprising?[1]

[1] Some will complain that I am here being too kind to the 'correspondence' theory: that I am appearing to concede its main claims. Nothing that I say about its *figurative* value as applied to a limited range of sentences should, however, be taken as affecting in any way my earlier remarks about its *literal* falsity: 'structure', 'reference', 'correspondence', these terms are all used here figuratively.

The example of the 'correspondence' theory of truth confirms our suspicion that the answer to the question, 'What makes utterances "reasons" for a conclusion?', depends, and must depend, upon the circumstances, and upon the type of conclusion involved. No single answer, no verbal formula comprehensive and general enough to cover all cases, can be hoped for.

Still, although we do not share the philosophers' hope for a universal answer to the question, 'What is "reasoning"?', we have no cause to be discouraged by this example in our search for particular answers, applicable to limited ranges of utterances. Indeed, the example is heartening, since it shows how natural and intelligible the logical criteria appropriate to a particular kind of conclusion can appear, when one bears in mind the circumstances in which utterances of that kind find their primary use.

Let us return, in conclusion, to the example with which this discussion of the 'correspondence theory' began: there are two things we must now notice. First, though some descriptive sentences *are* in some ways like pictures, this resemblance is not essential. Indeed, there is no reason in the world why a pictogram of a cat on a mat should necessarily serve as a 'description' or as a 'report' of a state of affairs at all. A picture, when used for linguistic purposes, is just as much a conventional symbol as a sentence written in a more normal fashion. Suppose, for instance, that a man comes up to me and hands me a photograph of a cat on a mat, saying, 'This is to be treated as a sentence'. Must I necessarily know what to do with it? How am I to tell whether he is trying to give me a report, say, rather than asking if the cat is on the mat, exclaiming at the beauty of cats, or issuing the injunction, 'No cats are to be allowed on mats'? After all, the metal discs bearing the head-on silhouette of a car, erected at the ends of some streets on the Continent, mean 'Cars are not allowed down this street', not 'There is a car coming down this street'—and even when one knows that they are intended as an injunction of some kind, one has to learn that it is this one, and not, 'Cars *are* to go down this street'.

Secondly, notice why it is that all those words like 'structure' and 'correspondence' which we have been using in this discussion are *figurative*. The reason is, that there are not two independent methods of identifying the two 'structures'—in the sentence, and in the world—*before* comparing them. Nevertheless, knowing as we do on what occasions we can properly and correctly say 'The cat is on the mat', we can *afterwards* point to the cat and the mat and the words 'cat' and 'mat', or the equivalent pictograms, and give a limited sense to the notions of 'structure' and 'correspondence'.



## 6.6 *Playing with Words*

In the case of descriptive sentences, the problem why one particular criterion of truth is appropriate was solved when we examined the purposes for which descriptions are used. The same intimate connection, between the logic of a mode of reasoning and the activities in which the reasoning plays its primary part, can be strikingly illustrated with the help of an especially simple (though artificial) example—that of an activity in which the reasoning operates as near as may be *functionlessly*.

In my childhood, we used to play a game with the following rule: starting from the letter A, and working through the alphabet, one had to make up sentences of the form:

> I love my love with an A, because she is artful;
> I hate her with an A, because she is arch;
> I take her to the 'Anchor'
> And feed her on artichokes;
> She comes from Aberdeen,
> And her name is Agnes Anstruther.

The primitive mode of reasoning involved in this game is the example required.

According to the rule of the game, only some of the possible reasons one might give for 'loving-with-an-A', or 'hating-with-an-A' are to be accepted. That your love is 'bashful' or 'comic' is a bad reason to give, either for 'loving-her-with-an-A' or for 'hating-her-with-an-A'; that she is 'awkward' or 'ambidextrous' is a good reason for either. Good reasons and bad reasons, correct and incorrect inferences, sound and unsound arguments, all are decided in this case by the rule of the game.

The exact nature of the rules we employ is comparatively arbitrary. The game would be none the worse, if one changed

them so that different 'reasons' were 'good' and 'bad' respectively, and so that the form of sentence required became:

> I love my love with an A because she is bashful;
> I hate her with an A because she is comic;
> I take her to the 'Dog and Duck'
> And feed her on earwigs;
> She comes from Fotheringay,
> And her name is Gertrude Gibson.

Despite the fact that the logical properties of the concepts 'loving-with-an-A' and 'hating-with-an-A', and the reasons relevant to each, are now distinct, the game could go on just as well as before.

The two most striking things about this example—the arbitrariness of the logical criteria, and the fact that it is only a game—are clearly connected. It is just because, being a game, this use of speech is comparatively *pointless*, that the logical criteria can be what you please. Rules there must be, rules for telling acceptable rigmaroles from unacceptable ones and good reasons from bad, if there is to be a game at all; but it makes no difference in the long run what these rules are, and the choice is therefore arbitrary.

You may, of course, object that in such a game one is just *playing* with words; and that, in consequence, one can hardly call it a *use* of reasoning, or a *use* of speech, at all. And it must be confessed that it is at the best only a borderline instance of 'reasoning'. But this only reinforces my point in citing it. All our more typical modes of reasoning are far from pointless, and the rules for distinguishing 'good' reasoning from 'bad' are correspondingly far from arbitrary.

### 6.7 *The Versatility of Reason*

In the case of word-games, as of descriptions, the nature of the logical criteria we are to apply is best understood from a study of the activity—and especially the point of the activity—of which the type of speech forms a part. But the example of a word-game has an additional value: that of showing what extremes of variety our uses of speech (and 'reasoning') include and, by implication, what a travesty of our practice it is to argue as though all our utterances were intended as descriptions.



If it is not completely nonsensical to talk of 'good and bad reasons' and of 'valid and invalid inferences', even over a mere word-game, how much less can it be so over the arguments we use in more important fields—in mathematics, in science, in ethics, in aesthetics, in expressing our reactions to things, in explaining our motives, in giving commands, and in our thousand-and-one other ways of using speech! And since, when we employ many of these modes of reasoning, it would be far-fetched to say that we were 'describing' anything, it is not to be expected that the logical criteria appropriate to these modes will be those relevant to descriptions. Rather, we must expect that every mode of reasoning, every type of sentence, and (if one is particular) every single sentence will have its own logical criteria, to be discovered by examining its individual, peculiar uses.

Hence arise the dangers of dogmatic attempts to define 'reasoning' uniquely. Some philosophers (like Hume) limit the scope of 'reasoning' to mathematics and science. Some stigmatise as 'pseudo-concepts' all concepts which do not 'refer to' specific objects or physical processes. Some are obsessed with particular, limited facets of 'truth'. And some dismiss all utterances other than those expressing factual propositions, on the ground that one cannot establish them in the way in which one establishes a factual hypothesis (as if one would want to!).

They would do none of these things if they recognised the full variety of purposes for which speech is used. Speech is no single-purpose tool. It is, in fact, more like a Boy Scout's knife (an implement with two kinds of blade, a screw-driver, a corkscrew, a tin-and-bottle opener, a file, an awl, and even a thing for taking stones out of horses' hooves); and, further, it is one which we continually shape and modify, adding new devices (modes of reasoning, and types of concept) to perform new functions, and grinding old ones afresh, in the light of experience, so that they shall serve their old, familiar, well-tried purposes better.[1]

### 6.8 *A New Approach to Our Problem*

We have examined two simple types of utterance—'descriptive sentences', for which the 'correspondence' theory specifies

[1] Wittgenstein makes use of the image of a tool-box in this context, likening different types of concept to different types of tool.

the criteria of truth and falsity, and the rigmaroles used in 'word-games', for which the logical criteria are what you choose to make them. The way in which each utterance is used, we discovered, was to be understood fully only as part of a larger activity: as soon as this frame was borne in mind, the logical properties of the mode of reasoning were seen to be related directly to the function it performs, and this again to the purpose of the activity of which it is a part. The logic of the utterances, on one side, and the point of the activity with which they are bound up, on the other, are as intimate and inseparable as the two faces of a coin.

Whether the same kind of relation holds for more complex modes of reasoning, and more complex types of sentence, remains to be seen. Our success so far at any rate encourages us to hope that, although the search for a general, universal answer to the question, 'What is "reasoning"?', was a mistaken one, we may still find answers applicable to individual modes of reasoning; and in particular that, by looking in the right way at the circumstances and activities in which our ethical utterances play their part, we may come to see how the logical criteria applicable to them are generated. Our central question, therefore, now takes the form:

Can we discover, from our knowledge of the kinds of human situation and activity in which ethical sentences find their primary use, the kinds of thing that are relevant as arguments for one course of action or another?

Can we, in other words, give an account of our ethical mode of reasoning which will bring out its characteristic function—that is, its contribution to the activities in which it is used?

Before we attempt to give such an account, let us try out our method on a mode of reasoning which is at once as complex as the ethical mode and less contentious. There is more agreement (because there is less general concern) over the purpose of science, and over the logic of scientific explanations, than there is over ethics and ethical standards; and yet the arguments which are presented in science are at least as complicated as any which are relevant to a moral decision. In both, there is the same duality, of individual judgements and general principles. In both, one encounters a contrast between the 'appearance' and the 'reality' —the scientist distinguishing between the 'apparent' colour of the sun at evening and its 'real' colour (quite apart from atmospheric refraction and so on); the moralist distinguishing those things which are 'really' good, and those actions which are 'really' right, from those things which we simply like and those actions which we simply feel like doing. Let us spend a little time, therefore, investigating the relation between the reasoning involved in a typical scientific argument, and the features of the situation to which it applies; and see what bearing this has on the special function of scientific argument, and of science itself. After we have done this, we can return and consider,



(i) whether ethics is a science—that is, whether the results of this investigation are immediately applicable to ethics;

(ii) if not, how the function of ethics differs from that of science; and

(iii) what kinds of argument are, in consequence, relevant to moral decisions.

# 7

# EXPERIENCE AND EXPLANATION

'. . . How build, unbuild, contrive
To save appearances. . . .'
MILTON, *Paradise Lost*, VIII, 81

## 7.1 *The Desire for an Explanation*

Suppose that we are going for a walk with some friends, and that I am carrying a walking-stick with me. If I pass the stick round, everyone will agree with me that it is straight; it will both look straight and feel straight to all of us. However, if we come to a stream and I plunge it half-way in, there will not be the same certainty. If we run our hands down the stick, we shall not feel, at the surface of the water, any change in the direction in which the stick is pointing—it will still feel straight to us all. But, if we are standing around, and each gives his own account of how it looks to him, we shall differ in what we say. Some will say, 'It's bent to the left'; others, 'It's bent to the right'; and one or two, 'It's just foreshortened'.

On encountering this phenomenon, you may react in any of a number of different ways. Your reaction may be one of wonder: you may simply gaze at the stick, ask me to pull it out and put it in again several times, so that you can take in what happens, and say, 'Isn't that marvellous?' Your reaction may be one of admiration: you may be struck by the way in which the contours of the bank enhance the bending of the stick, ask me to keep it there, and say, 'Isn't that a picture?' You may feel indifferent, wonder why anyone should be interested in the phenomenon, and turn away, saying, 'Well, what of it?' Or you may be surprised, not having expected this to happen, and say, 'Isn't that strange?' What you go on to do and say, what questions you ask, what consequences you draw, what investigations you undertake; all depend on the nature of your reaction—wonder, admiration, surprise or whatever it may be. For the moment, let us concentrate on the last of these—surprise.



Suppose that, in your reaction, what dominates is surprise at the unexpectedness of the phenomenon. The thing which disconcerts you about the situation is the way in which the evidence of our senses, originally unequivocal and unanimous, has become ambiguous and conflicting. There are obvious conflicts of three kinds:

(i) between the reports of the same observer about the same property at different times—first he said, 'It's straight'; now he says, 'It's bent';

(ii) between the reports of different observers about the same property at the same time—some say, 'It's bent to the left'; others say, 'It's bent to the right'; others again, 'It's just foreshortened';

(iii) between the evidence of different senses about the same property at the same time—looking at it, you would say it was bent but, to feel it, you would say it was straight.

In consequence of these conflicts, you ask, 'What is *really* the case? Is it really bent, or not? If so, to the left or to the right? And, if not, why does it look as though it were?' And, in asking these questions, you begin to demand an *explanation* of the phenomenon.

## 7.2 *Explanation and Expectation*

If what impresses you about the phenomenon is its unexpectedness, I can do a number of things to help you, and to satisfy your demand for an 'explanation'. To begin with, I can show you that this is something that always happens in such circumstances; that it is not a peculiarity of this particular stick or this particular stream, but that any straight piece of wood, metal or other solid, plunged into any smooth, level stretch of water, in stream or pond, in tank or wash-basin, looks the same. I can point out that the amount through which it appears to bend in such circumstances depends solely on the angle at which it is inserted and the direction from which one looks at it, and I can give you an equation connecting the angle through which it appears to bend with the angle of insertion and the direction of viewing; given this formula (Snell's Law), you will be able to decide beforehand how it will look when inserted at a new angle, or viewed from a fresh direction. Then I can show you that the same kind of thing

happens when the stick is embedded in substances other than water—in petrol, alcohol, glass or ice—but that the amount of bending differs from substance to substance; and I can give you a table of constants, with the help of which you can extend Snell's Law to cover these substances too (thus introducing the concept of 'refractive index'—that property of a substance, constant over a wide range of conditions, upon which the degree of bending depends). Again, I can relate this phenomenon to others with which you are familiar. By talking about the way in which the 'refractive index' varies with the density, and so with the temperature of the substance, I can connect up the bending of the stick and the shimmering of objects seen through the air over a fire. By referring to the way in which, in some materials, the 'refractive index' depends upon the colour of the light passed, I can connect this phenomenon with the colouring which appears round the edges of objects viewed through a cheap magnifying-glass—but there is no need to go on.

All these things will help to make the phenomenon seem less surprising. All of them will help to satisfy your demand for an 'explanation'. All of them do so by giving an explanation of the same kind, an explanation taken from *physics*, the appropriate science.[1] And each explanation is designed to show that, from our experience of optical phenomena, the bending of the stick was 'to be expected'.

If such an 'explanation' does satisfy you, that is no accident. The situation we have been considering is typical of those in which a 'scientific explanation' is called for, and to deal with which science was developed. And it shows us what, if anything, we must take as the special function of scientific explanation: namely, to bring our past experience to bear upon our present and future expectations, in such a way as to 'save appearances'[2] and turn the unexpected, as far as possible, into the expected.

[1] I believe the facts which I have quoted to be more or less correct. Nevertheless, it does not matter particularly whether they are or no: all that is necessary is that my 'explanation' should sound plausible, and be correct in form. If it is also correct in substance, that will simply help to avoid irrelevant objections.

[2] Cf. Milton, *Paradise Lost*, bk. VIII; and recall the old Greek phrase, σῴζειν τὰ φαινόμενα, discussed by Burnet, *Early Greek Philosophy*, p. 28.



### 7.3 *The Scientific Limitations of Everyday Concepts*

But why do we need a special mode of reasoning in order to do this? To understand the reason why, we must examine the way in which scientific theories and scientific concepts develop out of everyday life, language and experience.

The demand for an explanation arose, in the example given, out of discrepancies between different judgements about the 'straightness' of the walking-stick; that is, about something which, in everyday life, we treat as 'a single property'. It arose, that is, because the criteria of identity—of what is and what is not the 'same' property—which we take for granted in everyday life and crystallise in our everyday speech, turned out to be inadequate in the new situation. Let us analyse the nature of this crisis more closely.

When, in our ordinary life and affairs, we talk about the properties of objects, we are accustomed to using a variety of tests for the presence of any particular property. In sorting out 'straight' objects from 'bent' ones, for example, we treat straightness and bentness in different ways on different occasions:

(i) As 'simple qualities'

(*a*) to be told by eye;

(*b*) to be told by touch.

(ii) As 'complex qualities'

(*a*) to be told by sighting along the object, and taking the path of a light-ray as the standard;

(*b*) to be told by measurement, by comparing the length of the object with the distance between its ends, and taking the shortest distance as the standard;

(*c*) to be told by putting the object up against a template, and taking the edge of the template as the standard.

And, as a matter of fact, there is a large range of circumstances over which these tests give indistinguishable results.

A surveyor, when using his theodolite, and a carpenter, lifting his work to his eye to check his planing, take it for granted that a ray of light is straight—that is to say, they take the path of a ray as their standard. The same men, when working out a trigonometrical problem or setting out a job, take the shortest distance between two points as their standard instead. Neither of them

runs into any trouble as a result of using two logically independent criteria. If I look down a road, and see it stretching out directly before me to the very place I want to reach, I can be confident that by following it I shall be taking the shortest route to my destination. If, from a pile of brushwood, we pick out those branches that look straightest, we shall, as a matter of fact, make the same selection as we should if we picked out those which feel straightest, or compared their lengths with the distances between their ends, or sighted along them.

The success and utility of our concept of 'straightness' depends upon facts such as these. If either we or our environment were sufficiently different, it would be impossible for us to employ our present, everyday distinction between the 'bent' and the 'straight', impossible for us even to learn it. And things might very well have been so.

You could not, of course, teach anyone this distinction if all the instances you gave him were of the 'stick-in-water' variety: you have to start him on unambiguous ones—on sticks in air, broken and unbroken, lines drawn with and without rulers, Roman roads and English lanes, and such-like things. Yet, if the optical properties of the atmosphere were only fuller of discontinuities (like that between air and water) or of fluctuations (like those in the air over a fire) all the available instances would be of the ambiguous kind. And, given no favourable instances, we should have no chance of learning, and so no use for the distinction.

Again, if everyone were colour-blind—that is, if no one were able to learn our present distinction between the 'green' and the 'red'—we should have no use for the two everyday concepts 'green' and 'red': we should have only one. And, if our vision were such that we could no more tell the 'straight' from the 'bent' by eye than a colour-blind man can the 'green' from the 'red'—if we were all, by present standards, 'shape-blind'—then the existing distinction between the 'straight' and the 'bent', which we now regard as obvious and natural, would be useless, recondite and artificial.

The eventual breakdown of our everyday concepts, and the consequent demand for an 'explanation', are, therefore, pretty well inevitable. As long as we use multiple criteria for 'properties', sometimes telling them in one way, sometimes in another (and there is every reason why, for ordinary purposes, we should go on doing so), we are likely, sooner or later, to encounter a situation in which they lead to conflicting judgements. Our ordinary ideas about the 'properties' of the world can only lead us to predict and anticipate the developments going on around us under a limited range of conditions. Something unexpected always happens, and we are liable to be disconcerted when it does. If we want to be able to predict these occurrences, and not to be surprised by them, we have to give up the everyday account. Instead, we must try to find a more reliable one—a 'scientific' one.



### 7.4 *The Development of Scientific Theories and Concepts.* (I)

As a start, we tighten up our criteria of 'straightness', 'redness', 'rectangularity'—treating all properties as 'complex qualities', and demanding more accurate measurements, more specific descriptions. When this fails, we discard the everyday concepts, and replace them by a new set of concepts, defined by reference to the theory adopted.

These new concepts are of two kinds, those which are refinements of everyday concepts ('scientific qualities', I have called them) and those which have no counterparts in everyday speech —for example, physical and chemical constants like 'refractive index'. In either case, the danger of conflicts of the kind I have described is deliberately made as small as possible: the criteria for any 'property' are, as compared with those for everyday concepts, uniquely defined and highly specific. For the scientist, except in special circumstances which we shall come to later,[1] 'straightness' is completely unambiguous. The look, the feel, the ray of light, the template, all are of secondary importance. For him, a line is only 'straight' if it is the shortest line joining its end-points.

This unique choice of criterion, this particular definition of the scientific quality 'really straight', has a further, incidental advantage, of the greatest historical importance. It allows us to carry over bodily into science all the resources and paraphernalia of Euclidean and Cartesian geometry—the whole method of representing 'position' by three numerical co-ordinates, and

[1] §§ 8.5 and 8.6 below.

other physical properties as mathematical functions of these numerical co-ordinates. By adopting this unique definition of 'straightness', we can (so to speak) 'gear in' the results of mathematical research to the body of physical theory.

In our chosen example, it is the discrepancies between different criteria of 'straightness' and 'bentness' which lead to the demand for an explanation: physics explains away the apparent bentness of the stick as an optical effect. But the fact that it explains it away does not mean that the discordant observation, the look of bentness, is regarded as less important, at this stage, than the concordant ones. When we are searching for relatively constant and consistent factors in the situation, and formulating a new hypothesis to take account of them, it is the apparently conflicting reports which occupy us: we do not ask any more questions about the concordant reports until the discordant ones have been accounted for. If, adopting the classical physicist's definition of 'straightness', we say that the stick in the stream is 'really straight', then its *feel* does not worry us: if it still feels straight, so much the better—so much the less to explain! On the other hand, the look of the stick does puzzle us. It does not satisfy us to say, 'Now and then my eyes mislead me'—we must know when and how and why: so it is to the theory of optics that we turn for an explanation.

The theory of optics starts off from data about artificially simplified situations; for instance, the results of experiments in which extra-narrow rays of light are passed in precisely known directions through accurately ground prisms or lenses of unusually homogeneous glass. On the basis of data of this kind, formulae (such as Snell's Law, to which I have referred) are suggested relating the amounts by which the light-rays are deflected when they strike different prisms and lenses in different directions. Now, since the purpose of these formulae is to provide a more reliable indication of what was and is to be expected than any that can be given in terms of our everyday notions, the first test they must pass is that of accounting for all trustworthy past observations in the relevant field of study. And, when I say 'account for', I mean that they must show that these observations were to be expected: in other words, the theory must relate all the past phenomena to the conditions of the experiments concerned in the way which was actually found. Further, although the physicists who developed the theory of optics may have started by experimenting with prisms and lenses, the results of their work must be applicable also to the familiar kind of phenomenon we have taken as our example.



A theory which accounts in this way for all the observations made so far takes us half-way to our provisional goal. But as our account becomes more accurate, it looks less at the present and past, and more to the future. Our next step is to make trial predictions from the theory, using the methods of deductive logic—and of mathematics, where these are available—as a guide. Our theory will perhaps lead us to predict that, when I put the same stick into my bath, it will look bent again, and that while it stays in the air, it will continue (other things being equal) to look straight. As long as we find the relevant predictions confirmed by experience, we retain the theory. If, however, we are led to predict that the stick will only look bent in running water, and this expectation is falsified, so much the worse for the theory: we modify it, or abandon it in favour of a new one.

### 7·5 *The Development of Scientific Theories and Concepts.* (II)

Now we are hard to satisfy, and we do not want there to be any unpleasant surprises; so we keep on predicting and checking, first within the original field of experience, later changing the initial conditions one by one so as to pass beyond it. Eventually, we start making predictions in a field covered by another theory. At this stage, one of a number of things may happen.

(I) Our theory may, with no more than trifling changes, lead to the same predictions as the established theory, and the mathematical parts of the two theories may turn out to be formally similar. In this happy circumstance, the two theories can at once be unified and their concepts compared—'Light and radio waves are fundamentally of the same nature', we may explain; 'they are simply two manifestations of the same fundamental mechanism'.

(II) There may be a conflict: our theory may break down in the new field, and so prove inferior to the established theory: or it may do better than the established theory. In either case, we try to change the concepts and calculus of the deficient theory so

as to make them fit: we dislike unexplained limitations on our theories as much as any other unexplained phenomena.

(III) We may, however, be unable to do either of these things. Each theory may explain well enough the phenomena in its original field and their fields of study may be clearly connected, but neither may be able to touch the phenomena in the other's field; and this deadlock may continue for some time. An example is the deadlock between the 'electro-magnetic' (wave) theory of light and the 'photon' (particle) theory during the first quarter of this century. The wave-theory explained the propagation of light, the particle-theory the interaction of light with matter, and both with elegance and success; but neither of them seemed to apply in any way to the subject-matter of the other. In such a case, the solution lies in the development of a comprehensive set of concepts and a mathematical calculus embracing those of the two seemingly irreconcilable theories as special cases. This was done, for the wave-particle conflict, by the development of 'quantum mechanics' in the years beginning 1925.

(IV) Lastly, the two theories may lead to the same predictions, and yet their concepts and mathematical methods be totally dissimilar. An example of this, though for historical reasons not a completely pure one, can be found in the contrast between 'classical' and 'relativistic' physics. At the turn of the century, physicists were faced with certain experimental results—apparent anomalies in the behaviour of light—which would clearly be explained only after severe changes in physical theory. Einstein pointed out two ways of solving the problem. First, in his *Special Theory of Relativity*, he went half-way towards explaining the recalcitrant phenomena, without abandoning the framework of classical physics; and there is no logical reason why his work should not have been completed along these lines. Before this was done, however, he pointed out himself that the results could be more elegantly accounted for otherwise; and, in his *General Theory of Relativity*, he introduced an entirely different kind of mathematical framework, employing 'non-Euclidean' instead of 'Euclidean' geometry, and involving (as we shall see)[1] a different scientific criterion of 'straightness'. In the fields to which it is best applied, this 'relativistic' physics

[1] §§ 8.5 and 8.6 below.



has great advantages. The fresh mathematical methods introduced with it, such as the 'tensor calculus', make it possible for Sir Arthur Eddington to quote an equation, and then remark, 'There are about 280 billion terms on the right, and we proceed to rearrange those which do not vanish'[1]—an undertaking from which one might excusably shrink if one had at one's command only the less wholesale and sophisticated mathematics of classical physical theory. At the same time, over the greater part of physics, these more elementary methods can be made to produce the correct predictions without resort to marathon algebraic feats; and, since they *are* more elementary, they are generally preferred. The decision whether to develop a particular branch of physical theory along 'Euclidean' or 'non-Euclidean' lines is made, in practice, on an estimate of the relative difficulty of the mathematics involved in either case.

In deciding what scientific theory to adopt, therefore, we apply not one but a number of tests. The initial, and most important test is that of *predictive reliability*: the theory must show us that all the observations we have made in the relevant field of study were such as might have been expected, and it must give us the power to predict correctly future observations in the same field. The next test is that of *coherence*: if there are two theories which cover the same field of study with equal predictive reliability, we choose the one which fits in best with the theories established in adjacent fields of study—preferably in such a way that we can unify the mathematical and conceptual equipment of the two theories, and go deductively from the one to the other by different routes without inconsistency. Finally, if we have to choose between two theories which are both reliable, but which belong to different bodies of theory—e.g. one 'Euclidean' and one 'non-Euclidean'—we apply the test of *convenience*: the theory which produces the results with less effort on our part is the one we prefer.

### 7.6 *The Scope of Scientific Explanation*

This will do as an abbreviated account of the way in which scientific theories and concepts develop, and of the things which,

[1] A. S. Eddington, *The Mathematical Theory of Relativity*, p. 108: comment following equation 48.41.

at different stages in this development, lead us to accept and reject them. In the light of this survey, we can see in what kinds of situation the different questions which naturally arise in science can be answered and so, by elimination, the limits to the scope of science.

(I) Recall the situation from which we started. If, in your surprise at the appearance of the stick, you ask me, 'Why does this happen?', 'What is the explanation of this phenomenon?', or 'How do you explain this?', I shall know what you want; and I shall give you an answer demonstrating that the phenomenon might, from our experience in similar situations, have been expected. Your inquiry about the stick is intelligible because physics deals with just such phenomena; and the question, 'What is the explanation of this?', has the same force as, 'How would a physicist have come to predict this?' And I shall continue to understand you, so long as your inquiry is about phenomena which, from our experience, it is some good hoping to predict—phenomena, that is, with which we have some method of dealing, about which we have some kind of theory.

On the other hand, if you ask for a 'scientific explanation' of something which, in our experience, there is just no reason to expect, I shall be at a loss to know how to answer you. If, for example, you are impressed by the fact that the three Jones children all learnt to stand (or all died) on their respective birthdays, you may feel that you want an 'explanation' of the fact. But this is not the kind of event which we could have hoped to predict (and we mark this by calling it a 'coincidence'). Nothing which we did could 'show that it might have been expected', and I cannot therefore employ the familiar techniques of science to satisfy you. In consequence, if you do insist on asking me 'how I explain it', all I can do is to shrug my shoulders and reply that I do not attempt to. And, if you protest that I must have *some* explanation (still meaning a 'scientific' one), that is your mistake; for there are some situations in which the demand for a scientific explanation is out of place.

(II) Another class of questions naturally arising in science consists of those like, 'Is this explanation correct?' and 'Which of these explanations is correct?' If you ask me, 'Is this explanation correct?', I shall take your question as an inquiry as to whether this explanation would lead us to predict correctly all the recorded observations in the relevant field, or about the phenomena it leads us to expect, and the ways in which we can test it experimentally. Likewise, I shall know how to deal with the question, 'Which of these explanations is correct?', when it is equivalent to

(i) 'Which of them leads to the more reliable predictions?', or

(ii) 'Given that they are both satisfactory experimentally, which of them fits better with the established theories in adjoining branches of science?', or

(iii) 'Given that they are predictively equivalent, and both form parts of coherent bodies of science, which adopts the more laborious approach?'

But, once I have given you all the evidence about prediction, coherence and convenience, and you have agreed to it, I shall no longer understand what you are after, if you still ask, 'Is this explanation *true*?' or 'But which of these explanations is the *true* one?' As scientific questions, these no longer arise. Science is an activity of such a kind that there is no room for them.

(III) A third request, natural for a scientist, is for a 'unified theory', a comprehensive explanation covering more than one branch of science. If two branches of science are adjacent, and the established theories employ logically comparable concepts, it is proper to wonder about the relation between them, and to look for a theory covering both fields and embracing both established theories. You may very naturally and intelligibly inquire about the relation between (say) the physical and geometrical theories of optics—theories which, at any rate partly, lead to predictions about the same phenomena—and it is a physicist's business to provide the answer.

On the other hand, if two theories have been developed to explain phenomena in widely separated fields, or if they employ concepts whose logical characteristics are greatly different from one another—for example, the kinetic theory of matter and the psychology of musical composition—there is no question of their predictions being relevant to one another since, experimentally, they have no common ground. In such a case, it is only asking for trouble to try and 'unify' the theories.

Take a familiar example. Despite all the successes of classical physics in its own field (the explanation of the properties of gases, and so on), it is right to insist, with J. W. N. Sullivan, that

> the extension of the theory to the whole of phenomena was little more than an idle speculation. To suppose, e.g. that the Ninth Symphony was produced by the random collisions of little hard particles was never more, from the experimental point of view, than a pleasing fancy.[1]

And indeed, from the logical point of view, it was hardly even comprehensible. I can understand what it is to explain the apparent bending of the stick by reference to the shimmering of a heat-haze, but what would it be to 'explain' the genesis of the Choral Symphony by reference to the expansion and contraction of hydrogen? Nothing which I said about the properties of gases could ever quite remove your amazement at Beethoven's genius, or prove to you that the composition of the Choral Symphony 'might have been expected'.

(IV) A fourth class of questions, which we encounter when connecting up theories in adjoining fields of study, consists of conceptual questions, such as 'What is the nature of light?' If you ask me about the nature of light, I can talk to you about the close relationship between the physics of light and of electro-magnetism. I can point out the important similarities between optical and electro-magnetic phenomena, and I can demonstrate the formal equivalence of the mathematical theories (to confine myself for the moment to pre-quantum physics). And, as long as there is a place for an answer of this sort, these conceptual questions are intelligible. But once all this has been done the question, 'What is light?', has no further scientific answer. You may still feel worried, and want to go on asking, 'Is light *really* no more than very-very-high-frequency radio waves?' But, since all that science can do to satisfy you has been done, surprise is no longer appropriate: scientifically, at any rate, your question no longer has any sense.

### 7.7 *The 'Justification' of Science*

An examination of the situations in which one first looks for a 'scientific explanation', and of the function of the explanation

[1] J. W. N. Sullivan, *The Bases of Modern Science* (Penguin Edition), p. 206.



in these situations, can give one, therefore, an understanding of the *logic* of science. In talking of 'logic' I am here including both

(i) the tests to be applied to a 'scientific explanation' before one decides whether to accept it as 'correct', reject it as 'incorrect', or suspend judgement upon it, and

(ii) the limits to be placed on the scope of science, from which one is to decide when something that looks like a 'scientific' assertion or question has become either nonsensical or non-scientific.

Furthermore, such an examination shows us how self-contained an activity science is; how buttressed, cross-beamed and supported by one another its members are; and how solidly the whole structure rests upon its foundations of human life and purposes. Indeed, so natural and inevitable does the logic of science come to appear that one cannot help being a little taken aback, when a philosopher comes along and asks us how we justify, not just a particular one, but *all* scientific explanations.[1] For, apart from a more detailed and accurate account than I have had room to give of the way in which science develops—an account which could only be of use to us in justifying *this* explanation, as opposed to *that* one, and not 'explanations in general'—what kind of answer can be needed?

It is clear from the start that there is no room *within* science for the philosopher's inquiry. As a scientist, I can understand the questions, 'Is *this* explanation correct?' and '*Which* of these explanations is correct?': each presents me with a genuine choice, and requires from me a genuine decision. But if, instead, I am asked, 'Can *any* scientific explanation be correct?', what is wanted is completely mysterious: there are now no genuine alternatives, and so no choice or decision to be made. All that I can do is to explain the criteria which we do have for deciding whether or not to accept a scientific explanation, and point out how natural it is (bearing in mind the origins of science) that these should be our criteria.

[1] I am thinking in particular of the 'problem' of induction, as raised, for example, by Bertrand Russell in *The Problems of Philosophy* (1912), pp. 93 ff. It is interesting to reflect on Russell's desire for a 'justification' of science, in view of his parallel conviction that reason has no place in ethics: cf. footnote to § 11.10 below. One consideration which helps us to understand both of these views is this: it is possible, with the help of a 'Principle of Limited Variety' or such-like, to produce a demonstrative argument, which presents quite a vivid caricature of simple science, but there is no chance of doing the same for evaluation.

And it is not only within science that the philosopher's question is an odd one. For suppose that he explains what kind of 'justification' it is that he is asking for: then, either the results of such a 'justification' will be consistent with the conditions which we at present require a scientific argument to meet, or they will run counter to these conditions. Consider first the former case: what can his 'justification' do? It can only pay exaggerated attention to one feature of the situation at the expense of the others (in the same kind of way as the 'correspondence' and 'coherence' theories of truth), or else (like the 'pragmatic' theory) underline the fact that the process as a whole is one that pays.[1] (Needless to say, neither of these things is what the philosopher sees himself as doing.)

On the other hand, what are we to say if his results run counter to our present criteria? Shall we be asked to conclude that a scientist who has a reliable, coherent and convenient theory is nevertheless at fault in adopting it? There is something highly ridiculous and paradoxical in such a demand. Yet let us try to suppose that we do adopt the philosopher's new criteria in place of our present ones. Even if all that he does is to abolish some of our present tests as superfluous, we shall be in Queer Street.

Suppose, for instance, that he makes predictive reliability the only test that matters. Then, when faced with questions about why the Jones family all died on their birthdays, we can no longer excuse ourselves from taking any interest on the grounds—perfectly good grounds before we changed our criteria—that we are scientists, and that nothing in the results of science could have led us to predict that this would happen to them. We must now go in for fortune-telling and clairvoyance as well for, although they may be no more successful than science in this field, they can certainly be no less. And once we set up as clairvoyants there will be the same difficulty about calling us 'scientists' as there is about calling God a 'mathematician'[2] or a calculating-

[1] Cf. J. L. Austin on 'Other Minds', *Aristotelian Soc. Supp. Vol.* xx (1946), summed up by the author thus: 'Believing others, as it occurs in communicating, is one of the things we do, like giving promises or making inductions or playing games. If we press for an ultimate "justification" we shall only succeed in reducing it to something other than it is, or in proving that it pays.'

[2] Cf. Ayer, *Language, Truth and Logic* (2nd ed.), pp. 85–6.



boy an 'arithmetician':[1] going through all the proper steps (whether inductive or deductive) is something essential to genuine 'science' or 'mathematics'—jumping to the conclusion, even correctly, will not do.

The question, what makes a reason a 'good' reason in science, and what makes an argument or explanation a 'valid' one, can only be answered in terms of the reasons, arguments and explanations we *do* accept—namely, those which are predictively reliable, coherent and convenient. If we give up these criteria for others, we change the nature of our activity and, whatever we are now doing, it is no longer 'science'. The logical criteria applicable to scientific explanations are, in this respect, quite as intimately connected with the nature of the activity which we *call* 'science' as the logic and the activity of 'describing things' and the logic and the activity of 'loving-with-an-A'.

One possibility remains. Perhaps the philosopher is calling upon us to *give up* science. But, if this is the case, where are we to draw the line? In giving up science, are we to give up the activity in which natural science finds its use, for which science is our most highly developed tool? Are we, that is, to stop basing our expectations upon past experience?

These are not real questions, since we could not do so, even if we tried. The habit of basing our expectations upon experience of the past is so ingrained in us that it could only be suspended by a sustained effort of will—which itself could only be achieved by reference to experience of our own behaviour in the past. And, even if, by some miracle or misfortune, we were to lose the habit involuntarily, we should not do so for long, since the results would be quick, violent and fatal.

[1] Let me quote a paragraph from *The Times* for 17 December 1947:

'The *Cape Argus* reports a remarkable series of tests carried out by the professor of applied mathematics at Stellenbosch University on a broker's clerk who is reputed to have remarkable powers of mental calculation.

'In the presence of mathematical witnesses, and with careful safeguards, the young man was asked to answer the following questions: What is 58 times 73 times 67? What is 734 squared? What is 89 cubed? Multiply 961 by 579. Find the cube root of 84,567. He gave all the answers correctly in 39 seconds....

'Efforts are being made to enable the young man, who is one of a family of eight, to go to the university. *When at school his marks for mathematics were never high, because, although his answers were always right, his method of solving problems was always either unexplained or unorthodox.*' (My italics.)