

INTENTION

1. Very often, when a man says 'I am going to do such-and-such', we should say that this was an expression of intention. We also sometimes speak of an action as intentional, and we may also ask with what intention the thing was done. In each case we employ a concept of 'intention'; now if we set out to describe this concept, and took only one of these three kinds of statement as containing our whole topic, we might very likely say things about what 'intention' means which it would be false to say in one of the other cases. For example, we might say 'Intention always concerns the future'. But an action can be intentional without being concerned with the future in any way. Realising this might lead us to say that there are various senses of 'intention', and perhaps that it is thoroughly misleading that the word 'intentional' should be connected with the word 'intention', for an action can be intentional without having any intention in it. Or alternatively we may be tempted to think that only actions done with certain further intentions ought to be called intentional. And we may be inclined to say that 'intention' has a different sense when we speak of a man's intentions *simpliciter*—i.e. what he intends to do—and of his intention *in* doing or proposing something—what he aims at in it. But in fact it is implausible to say that the word is equivocal as it occurs in these different cases.

Where we are tempted to speak of 'different senses' of a word which is clearly not equivocal, we may infer that we are in fact pretty much in the dark about the character of the concept which it represents. There is, however, nothing wrong with taking a topic piecemeal. I shall therefore begin my enquiry by considering expressions of intention.

2. The distinction between an expression of intention and a prediction is generally appealed to as something intuitively clear. 'I am going to be sick' is usually a prediction; 'I am going to take a walk' usually an expression of intention. The distinction intended *is* intuitively clear, in the following sense: if

I say 'I am going to fail in this exam.' and someone says 'Surely you aren't as bad at the subject as that', I may make my meaning clear by explaining that I was expressing an intention, not giving an estimate of my chances.

If, however, we ask in philosophy what the difference is between e.g. 'I am going to be sick' as it would most usually be said, and 'I am going to take a walk', as *it* would most usually be said, it is not illuminating to be told that one is a prediction and the other the expression of an intention. For we are really asking what each of these is. Suppose it is said 'A prediction is a statement about the future'. This suggests that an expression of intention is not. It is perhaps the description—or expression—of a present state of mind, a state which has the properties that characterise it as an intention. Presumably what these are has yet to be discovered. But then it becomes difficult to see why they should be essentially connected with the future, as the intention seems to be. No one is likely to believe that it is an accident, a mere fact of psychology, that those states of mind which are intentions always have to do with the future, in the way that it is a fact of racial psychology, as one might say, that most of the earliest historical traditions concern heroic figures. And if you try to make being concerned with the future into a defining property of intentions, you can be asked what serves to distinguish this concern with the future from the predictive concern.

Let us then try to give some account of prediction. The following seems promising: a man says something with one inflection of the verb in his sentence; later that same thing, only with a changed inflection of the verb, can be called true (or false) in face of what has happened later.

Now by this criterion, commands and expressions of intention will also be predictions. In view of the difficulties described above, this may not constitute an objection. Adopting a hint from Wittgenstein (*Philosophical Investigations* §§ 629–30) we might then first define prediction in general in some such fashion, and then, among predictions, distinguish between commands, expressions of intention, estimates, pure prophecies, etc. The 'intuitively clear' distinction we spoke of turns out to be a distinction between expressions of intention and estimates. But

a single utterance may function as more than one of these kinds of prediction. E.g. when a doctor says to a patient in the presence of a nurse 'Nurse will take you to the operating theatre', this may function both as an expression of his intention (if it is in it that his decision as to what shall happen gets expressed) and as an order, as well as being information to the patient; and it is this latter in spite of being in no sense an estimate of the future founded on evidence, nor yet a guess or prophecy; nor does the patient normally *infer* the information from the fact that the doctor said that; he would say that the doctor *told* him. This example shews that the indicative (descriptive, informatory) character is not the distinctive mark of 'predictions' *as opposed to* 'expressions of intention', as we might at first sight have been tempted to think.

An imperative will be a description of some future action, addressed to the prospective agent, and cast in a form whose point in the language is to make the person do what is described. I say that this is its point in the language, rather than that it is the purpose of the speaker, partly because the speaker might of course give an order with some purpose quite other than that it should be executed (e.g. so that it should *not* be executed), without detriment to its being an order.

Execution-conditions for commands correspond to truth-conditions for propositions. What are the reasons other than a dispensable usage for not calling commands true and false according as they are obeyed or disobeyed?

An order will usually be given with some intention or other, but is not as such the expression of a volition; it is simply a description of an action cast in a special form; this form is sometimes a special inflection and sometimes a future tense which has other uses as well.

Orders are usually criticised for being sound or unsound rather than for being fulfilled or not fulfilled; but this does not serve to distinguish orders from estimates of the future, since the same may hold for estimates of the future, where these are scientific. (Unscientific estimates are of course praised for being fulfilled rather than for being well-founded, as no one knows what a good foundation is for an unscientific estimate—e.g. a political one.) But there is a difference between the types of

ground on which we call an order, and an estimate of the future, sound. The reasons justifying an order are not ones suggesting what is probable, or likely to happen, but e.g. ones suggesting what it would be good to make happen with a view to an objective, or with a view to a sound objective. In this regard, commands and expressions of intention are similar.

It is natural to feel an objection both to calling commands, and to calling expressions of intention, predictions. In the case of commands, the reason lies in the superficial grammar, and just because of this is more easily disposed of. In the case of intentions, superficial grammar would rather incline us to accept the diagnosis, since a common form of expression of intention is a simple future tense, and indeed, this use of the future tense must play a dominant part in any child's learning of it. But our objections are deeper rooted.

If I do not do what I said I would, I am not supposed to have made a mistake, or even necessarily to have lied; so it seems that the truth of a statement of intention is not a matter of my doing what I said. But why should we not say: this only shows that there are other ways of saying what is not true, besides lying and being mistaken?

A lie, however, is possible here; and if I lie, what I say is a lie because of something present, not future. I might even be lying in saying I was going to do something, though I afterwards did it. The answer to this is that a lie is an utterance contrary to one's mind, and one's mind may be either an opinion, or a mind to make something the case. That a lie is an utterance contrary to one's mind does not mean that it is a false report of the contents of one's mind, as when one lies in response to the query 'A penny for your thoughts'.

One might not have a 'mind' to do something, distinguishable from uttering the words. And then, as Quine once put it (at a philosophical meeting), one might do the thing 'to make an honest proposition' of what one had said. For if I don't do what I said, what I said was not true (though there might not be a question of my *truthfulness* in saying it). But the reason why Quine's remark is a joke is that this falsehood does not necessarily impugn what I *said*. In some cases the facts are, so to speak, impugned for not being in accordance with the words,

rather than *vice versa*. This is sometimes so when I change my mind; but another case of it occurs when e.g. I write something other than I think I am writing: as Theophrastus says (*Magna Moralia*,¹ 1189b 22), the mistake here is one of performance, not of judgment. There are other cases too: for example, St. Peter did not *change his mind* about denying Christ; and yet it would not be correct to say he made a lying promise of faithfulness.

A command is essentially a sign (or symbol), whereas an intention can exist without a symbol; hence we speak of commands, not of the expression of commanding; but of the *expression of intention*. This is another reason for the very natural idea that in order to understand the expression of intention, we ought to consider something internal, i.e. what it is an expression of. This consideration disinclines us to call it a prediction—i.e. a description of something future. Even though that is just what 'I'll do such-and-such' actually looks like, and even though 'I intend to go for a walk but shall not go for a walk' does sound in some way contradictory.

Intention appears to be something that we can express, but which brutes (which e.g. do not give orders) can *have*, though lacking any distinct expression of intention. For a cat's movements in stalking a bird are hardly to be called an expression of intention. One might as well call a car's stalling the *expression* of its being about to stop. Intention is unlike emotion in this respect, that the expression of it is purely conventional; we might say 'linguistic', if we will allow certain bodily movements with a conventional meaning to be included in language. Wittgenstein seems to me to have gone wrong in speaking of the 'natural expression of an intention' (*Philosophical Investigations* § 647).

3. We need a more fruitful line of enquiry than that of considering the verbal expression of intention, or of trying to consider what it is an expression of. For if we consider just the verbal expression of intention, we arrive only at its being a—queer—species of prediction; and if we try to look for what it is an expression of, we are likely to find ourselves in one or other of several dead ends, e.g.: psychological jargon about 'drives' and

¹ Assuming that we are correctly told that Theophrastus was the author.

'sets'; reduction of intention to a species of desire, i.e. a kind of emotion; or irreducible intuition of the meaning of 'I intend'.

Looking at the verbal expression of intention is indeed of use for avoiding these particular dead-ends. They are all reached in consequence of leaving the distinction between estimation of the future and expression of intention as something that just is intuitively obvious. A man says 'I am going for a walk' and we say 'that is an expression of intention, not a prediction'. But how do we know? If we asked him, no doubt he would tell us; but what does he know, and how? Wittgenstein has shown the impossibility of answering this question by saying 'He recognizes himself as having, or as having had, an intention of going for a walk, or as having meant the words as an expression of intention'. If this were correct, there would have to be room for the possibility that he misrecognizes. Further, when we remember having meant to do something, what memory reveals as having gone on in our consciousness is a few scanty items at most, which by no means add up to such an intention; or it simply prompts us to use the words 'I meant to . . .', without even a mental picture of which we judge the words to be an appropriate description. The distinction, then, cannot be left to be intuitively obvious, except where it is used to answer the question in what sense a man meant the form of words 'I am going to . . .' on a particular occasion.

We might attempt to make the distinction out by saying: an expression of intention is a description of something future in which the speaker is some sort of agent, which description he justifies (if he does justify it) by reasons for acting, sc. reasons why it would be useful or attractive if the description came true, not by evidence that it is true. But having got so far, I can see nowhere else to go along this line, and the topic remains rather mystifying. I once saw some notes on a lecture of Wittgenstein in which he imagined some leaves blown about by the wind and saying 'Now I'll go this way . . . now I'll go that way' as the wind blew them. The analogy is unsatisfactory in apparently assigning no role to these predictions other than that of an unnecessary accompaniment to the movements of the leaves. But it might be replied: what do you mean by an 'unnecessary' accompaniment? If you mean one in the absence of which the

movements of the leaves would have been just the same, the analogy is certainly bad. But how do you know what the movements of the leaves would have been if they had not been accompanied by those thoughts? If you mean that you could calculate their movements just by knowing the speed and direction of the winds and the weight and other properties of the leaves, are you insisting that such calculations could not include calculations of their thoughts?—Wittgenstein was discussing free will when he produced this analogy; now the objection to it is not that it assigns a false role to our intentions, but only that it does not describe their role at all; this, however, was not its purpose. That purpose was clearly *some* denial of free will, whether we take the wind as a symbol for the physical forces that affect us, or for God or fate. Now it may be that a correct description of the role of intention in our actions will not be relevant to the question of free will; in any case I suspect that this was Wittgenstein's view; therefore in giving this anti-freewill picture he was at liberty simply to leave the role of intention quite obscure.

Now our account of expressions of intention, whereby they are distinguished from estimates of the future, leaves one in very much the same position as does the picture of the wind blowing the leaves. People do in fact give accounts of future events in which they are some sort of agents; they do not justify these accounts by producing reasons why they should be believed but, if at all, by a different sort of reason; and these accounts are very often correct. This sort of account is called an expression of intention. It just does occur in human language. If the concept of 'intention' is one's quarry, this enquiry has produced results which are indeed not false but rather mystifying. What is meant by 'reason' here is obviously a fruitful line of enquiry; but I prefer to consider this first in connexion with the notion of intentional action.

4. I therefore turn to a new line of enquiry: how do we tell someone's intentions? or: what kind of true statements about people's intentions can we certainly make, and how do we know that they are true? That is to say, is it possible to find types of statement of the form 'A intends X' which we can say have a

great deal of certainty? Well, if you want to say at least some true things about a man's intentions, you will have a strong chance of success if you mention what he actually did or is doing. For whatever else he may intend, or whatever may be his intentions in doing what he does, the greater number of the things which you would say straight off a man did or was doing, will be things he intends.

I am referring to the sort of things you would say in a law court if you were a witness and were asked what a man was doing when you saw him. That is to say, in a very large number of cases, your selection from the immense variety of true statements about him which you might make would coincide with what he could say he was doing, perhaps even without reflection, certainly without adverting to observation. I am sitting in a chair writing, and anyone grown to the age of reason in the same world would know this as soon as he saw me, and in general it would be his first account of what I was doing; if this were something he arrived at with difficulty, and what he knew straight off were precisely how I was affecting the acoustic properties of the room (to me a very recondite piece of information), then communication between us would be rather severely impaired.

In this way, with a view to shewing roughly the range of things to be discovered here, I can take a short cut here, and discuss neither how I am to select from the large number of true statements I could make about a person, nor what is involved in the existence of such a straight-off description as 'She is sitting in a chair and writing'. (Not that this does not raise very interesting questions. See *Philosophical Investigations*, p. 59, (b): 'I see a picture: it shows a man leaning on a stick and going up a steep path. How come? Couldn't it look like that if he were sliding downhill in that position? Perhaps a Martian would give that description.' *Et passim*.) All I am here concerned to do is note the fact: we can simply say 'Look at a man and say what he is doing'—i.e. say what would immediately come to your mind as a report to give someone who could not see him and who wanted to know what was to be seen in that place. In most cases what you will say is that the man himself knows; and again in most, though indeed in fewer, cases you will be reporting not merely what he is doing, but *an* intention of his—namely, to do that

thing. What is more, if it is not an intention of his, this will for the most part be clear without asking him.

Now it can easily seem that in general the question what a man's intentions are is only authoritatively settled by him. One reason for this is that in general we are interested, not just in a man's intention *of* doing what he does, but in his intention *in* doing it, and this can very often not be seen from seeing what he does. Another is that in general the question whether he intends to do what he does just does not arise (because the answer is obvious); while if it does arise, it is rather often settled by asking him. And, finally, a man can form an intention which he then does nothing to carry out, either because he is prevented or because he changes his mind: but the intention itself can be complete, although it remains a purely interior thing. All this conspires to make us think that if we want to know a man's intentions it is into the contents of his mind, and only into these, that we must enquire; and hence, that if we wish to understand what intention is, we must be investigating something whose existence is purely in the sphere of the mind; and that although intention issues in actions, and the way this happens also presents interesting questions, still what physically takes place, i.e. what a man actually does, is the very last thing we need consider in our enquiry. Whereas I wish to say that it is the first. With this preamble to go on to the second head of the division that I made in § 1: intentional action.

5. What distinguishes actions which are intentional from those which are not? The answer that I shall suggest is that they are the actions to which a certain sense of the question 'Why?' is given application; the sense is of course that in which the answer, if positive, gives a reason for acting. But this is not a sufficient statement, because the question "What is the relevant sense of the question 'Why?'" and "What is meant by 'reason for acting'?" are one and the same.

To see the difficulties here, consider the question, 'Why did you knock the cup off the table?' answered by 'I thought I saw a face at the window and it made me jump'. Now, so far I have only characterised reason for acting by opposing it to evidence for supposing the thing will take place—but the 'reason'

here was not evidence that I was going to knock the cup off the table. Nor can we say that since it mentions something previous to the action, this will be a cause rather than a reason; for if you ask 'Why did you kill him?' the answer 'He killed my father' is surely a reason rather than a cause, but what it mentions is previous to the action. It is true that we don't ordinarily think of a case like giving a sudden start when we speak of a reason for acting. "Giving a sudden start", someone might say, "is not *acting* in the sense suggested by the expression 'reason for acting'. Hence, though indeed we readily say e.g. 'What was the reason for your starting so violently?' this is totally unlike 'What is your reason for excluding so-and-so from your will?' or 'What is your reason for sending for a taxi?'" But what *is* the difference? In neither case is the answer a piece of evidence. Why is giving a start or gasp not an 'action', while sending for a taxi, or crossing the road, is one? The answer cannot be "Because the answer to the question 'why?' may give a *reason* in the latter cases", for the answer may 'give a reason' in the former cases too; and we cannot say "Ah, but not a reason for *acting*"; we should be going round in circles. We need to find the difference between the two kinds of 'reason' without talking about 'acting'; and if we do, perhaps we shall discover what is meant by 'acting' when it is said with this special emphasis.

It will hardly be enlightening to say: in the case of the sudden start the 'reason' is a *cause*; the topic of causality is in a state of too great confusion; all we know is that this is one of the places where we do use the word 'cause'. But we also know that this is a rather strange case of causality; the subject is able to give the cause of a thought or feeling or bodily movement in the same kind of way as he is able to state the place of his pain or the position of his limbs.

Nor can we say: "—Well, the 'reason' for a movement is a cause, and not a reason in the sense of 'reason for acting', when the movement is involuntary; it is a reason, as opposed to a cause, when the movement is voluntary and intentional." This is partly because in any case the object of the whole enquiry is really to delineate such concepts as the voluntary and the intentional, and partly because one can also give a 'reason' which is only a 'cause' for what is voluntary and intentional. E.g. "Why

are you walking up and down like that?"—"It's that military band; it excites me". Or "What made you sign the document at last?"—"The thought: 'It is my duty' kept hammering away in my mind until I said to myself 'I can do no other', and so signed."

It is very usual to hear that such-and-such are what we *call* 'reasons for acting' and that it is 'rational' or 'what we *call* rational' to act for reasons; but these remarks are usually more than half moralistic in meaning (and moralism, as Bradley remarked, is bad for thinking); and for the rest they leave our conceptual problems untouched, while pretending to give a quick account. In any case, this pretence is not even plausible, since such remarks contain no hint of what it is to act for reasons.

6. To clarify the proposed account, "Intentional actions are ones to which a certain sense of the question 'why?' has application", I will both explain this sense and describe cases shewing the question *not* to have application. I will do the second job in two stages because what I say in the first stage of it will be of use in helping to explain the relevant sense of the question 'why?'.

This question is refused application by the answer: 'I was not aware I was doing that'. Such an answer is, not indeed a proof (since it may be a lie), but a claim, that the question 'Why did you do it (are you doing it)?', in the required sense, has no application. It cannot be plausibly given in every case; for example, if you saw a man sawing a plank and asked 'Why are you sawing that plank?', and he replied 'I didn't know I was sawing a plank', you would have to cast about for what he might mean. Possibly he did not know the word 'plank' before, and chooses this way of expressing that. But this question as to what he might mean need not arise at all—e.g. if you ask someone why he is standing on a hose-pipe and he says 'I didn't know I was'.

Since a single action can have many different descriptions, e.g. 'sawing a plank', 'sawing oak', 'sawing one of Smith's planks', 'making a squeaky noise with the saw', 'making a great deal of sawdust' and so on and so on, it is important to notice that a man may know that he is doing a thing under one description, and not under another. Not every case of this is a

case of his knowing that he is doing one part of what he is doing and not another (e.g. he knows he is sawing but not that he is making a squeaky noise with the saw). He may know that he is sawing a plank, but not that he is sawing an oak plank or Smith's plank; but sawing an oak plank or Smith's plank is not something else that he is doing besides just sawing the plank that he is sawing. For this reason, the statement that a man knows he is doing X does not imply the statement that, concerning anything which is also his doing X, he knows that he is doing that thing. So to say that a man knows he is doing X is to give a description of what he is doing *under which* he knows it. Thus, when a man says 'I was not aware that I was doing X', and so claims that the question 'Why?' has no application, he cannot always be confuted by the fact that he was attentive to those of his own proceedings in which doing X consisted.

7. It is also clear that one is refusing application to the question 'Why?' (in the relevant sense) if one says: 'It was involuntary', even though the action was something of which one was aware. But I cannot use this as it stands, since the notion of the involuntary pretty obviously covers notions of exactly the type that a philosophical enquiry into intention ought to be elucidating.

Here, digressing for a moment, I should like to reject a fashionable view of the terms 'voluntary' and 'involuntary', which says they are appropriately used only when a person has done something untoward. If anyone is tempted by this view, he should consider that physiologists are interested in voluntary action, and that they are not giving a special technical sense to the word. If you ask them what their criterion is, they say that if they are dealing with a grown human they ask him, and if with an animal, they take movements in which the animal is e.g. trying to get at something, say food. That is, the movement by which a dog cocked its ear at a sudden sound would not be used as an example.

This does not mean that every description of action in which its voluntariness can be considered is of interest to physiologists. Of course they are only interested in bodily movements.

We can also easily get confused by the fact that 'involuntary'

neither means simply non-voluntary, nor has an unproblematic sense of its own. In fact this pair of concepts is altogether very confusing. Consider the four following examples of the involuntary:

(a) The peristaltic movement of the gut.

(b) The odd sort of jerk or jump that one's whole body sometimes gives when one is falling asleep.

(c) 'He withdrew his hand in a movement of involuntary recoil.'

(d) 'The involuntary benefit I did him by a stroke I meant to harm him.'

Faced with examples like (c) and (d), how can I introduce 'It was involuntary' as a form for rejecting the question 'Why?' in the special sense which I want to elucidate—when the whole purpose of the elucidation is to give an account of the concept 'intentional'? Obviously I cannot. There is however a class of the things that fall under the concept 'involuntary', which it is possible to introduce without begging any questions or assuming that we understand notions of the very type I am professing to investigate. Example (b) belongs to this class, which is a class of bodily movements in a purely physical description. Other examples are tics, reflex kicks from the knee, the lift of the arm from one's side after one has leaned heavily with it up against a wall.

8. What is required is to describe this class without using any notions like 'intended' or 'willed' or 'voluntary' and 'involuntary'. This can be done as follows: we first point out a particular class of things which are true of a man: namely the class of things which he *knows without observation*. E.g. a man usually knows the position of his limbs without observation. It is without observation, because nothing *shows* him the position of his limbs; it is not as if he were going by a tingle in his knee, which is the sign that it is bent and not straight. Where we can speak of separately describable sensations, having which is in some sense our criterion for saying something, then we can speak of observing that thing; but that is not generally so when we know the position of our limbs. Yet, without prompting, we *can say* it. I say however that we *know* it and not merely *can say*

it, because there is a possibility of being right or wrong: there is point in speaking of knowledge only where a contrast exists between 'he *knows*' and 'he (merely) *thinks* he knows'. Thus, although there is a similarity between giving the position of one's limbs and giving the place of one's pain, I should wish to say that one ordinarily *knows* the position of one's limbs, without observation, but not that being able to say where one feels pain is a case of something known. This is not because the place of pain (the feeling, not the damage) has to be accepted by someone I tell it to; for we can imagine circumstances in which it is not accepted. As e.g. if you say that your foot, not your hand, is very sore, but it is your hand you nurse, and you have no fear of or objection to an inconsiderate handling of your foot, and yet you point to your foot as the sore part: and so on. But here we should say that it was difficult to guess what you could mean. Whereas if someone says that his leg is bent when it is straight, this may be surprising but is not particularly obscure. He is wrong in what he says, but not unintelligible. So I call this sort of being able to say 'knowledge' and not *merely* 'being able to say'.

Now the class of things known without observation is of general interest to our enquiry because the class of intentional actions is a sub-class of it. I have already said that 'I was not aware I was doing that' is a rejection of the question 'Why?' whose sense we are trying to get at; here I can further say 'I knew I was doing that, but only because I observed it' would also be a rejection of it. E.g. if one noticed that one operated the traffic lights in crossing a road.

But the class of things known without observation is also of special interest in this part of our enquiry, because it makes it possible to describe the particular class of 'involuntary actions' which I have so far indicated just by giving a few examples: these are actions like the example (b) above, and our task is to mark off this class without begging the questions we are trying to answer. Bodily movements like the peristaltic movement of the gut are involuntary; but these do not interest us, for a man does not know his body is making them except by observation, inference, etc. The involuntary that interests us is restricted to the class of things known without observation; as you would

know even with your eyes shut that you had kicked when the doctor tapped your knee, but cannot identify a sensation by which you know it. If you speak of 'that sensation which one has in reflex kicking, when one's knee is tapped', this is not like e.g. 'the sensation of going down in a lift'. For though one might say 'I thought I had given a reflex kick, when I hadn't moved' one would never say e.g. 'Being told startling news gives one that sensation': the sensation is not separable, as the sensation 'like going down in a lift' is.

Now among things known without observation must be included the causes of some movements. E.g. 'Why did you jump back suddenly like that?' 'The leap and loud bark of that crocodile made me jump'. (I am not saying I did not observe the crocodile barking; but I did not observe that making me jump.) But in examples like (b) the cause of motion is known *only* through observation.

This class of involuntary actions, then, is the class of movements of the body, in a purely physical description, which are known without observation, and where there is no such thing as a *cause* known without observation. (Thus my jump backwards at the leap and bark of the crocodile does *not* belong to this subclass of involuntary actions.) This subclass can be described without our first having clarified the concept 'involuntary'. To assign a movement to it will be to reject the question 'Why?'

9. I first, in considering expressions of intention, said that they were predictions justified, if at all, by a reason for acting, as opposed to a reason for thinking them true. So I here already distinguished a sense of 'Why?', in which the answer mentions evidence. 'There will be an eclipse tomorrow'.—'Why?' 'Because . . .'—and an answer is the reason for thinking so. Or 'There was an ancient British camp here'. 'Why?'—and an answer is the reason for thinking so. But as we have already noted, an answer to the question 'Why?' which does not give reason for thinking the thing true does not *therefore* give a reason for acting. It may mention a cause, and this is far from what we want. However we noticed that there are contexts in which there is some difficulty in describing the distinction between a

cause and a reason. As e.g. when we give a ready answer to the question 'Why did you knock the cup off the table?'—'I saw such-and-such and *it made me jump.*'

Now we can see that the cases where this difficulty arises are just those where the cause itself *qua* cause (or perhaps one should rather say: the causation itself) is in the class of things known without observation.

10. I will call the type of cause in question a '*mental cause*'. Mental causes are possible, not only for actions ('The martial music excites me, that is why I walk up and down') but also for feelings and even thoughts. In considering actions, it is important to distinguish between mental causes and motives; in considering feelings, such as fear or anger, it is important to distinguish between mental causes and objects of feeling. To see this, consider the following cases:

A child saw a bit of red stuff on a turn in a stairway and asked what it was. He thought his nurse told him it was a bit of Satan and felt dreadful fear of it. (No doubt she said it was a bit of satin.) What he was frightened of was the bit of stuff; the cause of his fright was his nurse's remark. The object of fear may be the cause of fear, but, as Wittgenstein¹ remarks, is not *as such* the cause of fear. (A hideous face appearing at the window would of course be both cause and object, and hence the two are easily confused). Or again, you may be angry *at* someone's action, when what *makes* you angry is some reminder of it, or someone's telling you of it.

This sort of cause of a feeling or reaction may be reported by the person himself, as well as recognised by someone else, even when it is not the same as the object. Note that this sort of causality or sense of 'causality' is so far from accommodating itself to Hume's explanations that people who believe that Hume pretty well dealt with the topic of causality would entirely leave it out of their calculations; if their attention were drawn to it they might insist that the word 'cause' was inappropriate or was quite equivocal. Or conceivably they might try to give a Humian account of the matter as far as concerned the outside observer's recognition of the cause; but hardly for the patient's.

¹ *Philosophical Investigations* § 476.

11. Now one might think that when the question 'Why?' is answered by giving the intention with which a person acts—for example by mentioning something future—this is also a case of a mental cause. For couldn't it be recast in the form: 'Because I wanted . . .' or 'Out of a desire that . . .'? If a feeling of desire to eat apples affects me and I get up and go to a cupboard where I think there are some, I might answer the question what led to this action by mentioning the desire as having made me . . . etc. But it is not in all cases that 'I did so and so in order to . . .' can be backed up by 'I *felt* a desire that . . .'. I may e.g. simply hear a knock on the door and go downstairs to open it without experiencing any such desire. Or suppose I feel an upsurge of spite against someone and destroy a message he has received so that he shall miss an appointment. If I describe this by saying 'I wanted to make him miss that appointment', this does not necessarily mean that I had the thought 'If I do this, he will . . .' and that affected me with a desire of bringing it about, which led up to my doing so. This may have happened, but need not. It could be that all that happened was this: I read the message, had the thought 'That unspeakable man!' with feelings of hatred, tore the message up, and laughed. Then if the question 'Why did you do that?' is put by someone who makes it clear that he wants me to mention the mental causes—e.g. what went on in my mind and issued in the action—I should perhaps give this account; but normally the reply would be no such thing. That particular enquiry is not very often made. Nor do I wish to say that it always has an answer in cases where it can be made. One might shrug or say 'I don't know that there was any definite history of the kind you mean', or 'It merely occurred to me. . . .'

A '*mental cause*', of course, need not be a mental event, i.e. a thought or feeling or image; it might be a knock on the door. But if it is not a mental event, it must be something perceived by the person affected—e.g. the knock on the door must be heard—so if in this sense anyone wishes to say it is always a mental event, I have no objection. A mental cause is what someone would describe if he were asked the specific question: what produced this action or thought or feeling on your part: what did you see or hear or feel, or what ideas or images cropped up in

your mind, and led up to it? I have isolated this notion of a mental cause because there is such a thing as this question with this sort of answer, and because I want to distinguish it from the ordinary senses of 'motive' and 'intention', rather than because it is in itself of very great importance; for I believe that it is of very little. But it is important to have a clear idea of it, partly because a very natural conception of 'motive' is that it is what *moves* (the very word suggests that)—glossed as 'what *causes*' a man's actions etc. And 'what causes' them is perhaps then thought of as an event that brings the effect about—though how it does—i.e. whether it should be thought of as a kind of pushing in another medium, or in some other way—is of course completely obscure.

12. In philosophy a distinction has sometimes been drawn between our motives and our intentions in acting as if they were quite different things. A man's intention is *what* he aims at or chooses; his motive is what determines the aim or choice; and I suppose that 'determines' must here be another word for 'causes'.

Popularly motive and intention are not treated as so distinct in meaning. E.g. we hear of 'the motive of gain'; some philosophers have wanted to say that such an expression must be elliptical; gain must be the *intention*, and *desire of gain* the motive. Asked for a motive, a man might say 'I wanted to . . .', which would please such philosophers; or 'I did it in order to . . .', which would not; and yet the meaning of the two phrases is here identical. When a man's motives are called good, this may be in no way distinct from calling his intentions good—e.g. he only wanted to make peace among his relations.

Nevertheless there is even popularly a distinction between the meaning of 'motive' and the meaning of 'intention'. E.g. if a man kills someone, he may be said to have done it out of love and pity, or to have done it out of hatred; these might indeed be cast in the forms 'to release him from this awful suffering', or 'to get rid of the swine'; but though these are forms of expression suggesting objectives, they are perhaps expressive of the spirit in which the man killed rather than descriptive of the end to which the killing was a means—a future state of affairs to be

produced by the killing. And this shows us part of the distinction that there is between the popular senses of motive and intention. We should say: popularly, 'motive for an action' has a rather wider and more diverse application than 'intention with which the action was done'.

When a man says what his motive was, speaking popularly, and in a sense in which 'motive' is not interchangeable with 'intention', he is not giving a 'mental cause' in the sense that I have given to that phrase.—The fact that the mental causes were such-and-such may indeed help to make his claim intelligible. And further, though he may say that his motive was this or that one straight off and without lying—i.e. without saying what he knows or even half knows to be untrue—yet a consideration of various things, which may include the mental causes, might possibly lead both him and other people to judge that his declaration of his own motive was false. But it appears to me that the mental causes are seldom more than a very trivial item among the things that it would be reasonable to consider. As for the importance of considering the motives of an action, as opposed to considering the intention, I am very glad not to be writing either ethics or literary criticism, to which this question belongs.

Motives may explain actions to us; but that is not to say that they 'determine', in the sense of causing, actions. We do say: 'His love of truth caused him to . . .' and similar things, and no doubt such expressions help us to think that a motive must be what produces or brings about a choice. But this means rather 'He did this in that he loved the truth'; it interprets his action.

Someone who sees the confusions involved in radically distinguishing between motives and intentions and in defining motives, so distinct, as the determinants of choice, may easily be inclined to deny both that there is any such thing as mental causality, and that 'motive' means anything but intention. But both of these inclinations are mistaken. We shall create confusion if we do not notice (*a*) that phenomena deserving the name of mental causality exist, for we can make the question 'Why?' into a request for the sort of answer that I considered under that head; (*b*) that mental causality is not restricted to choices or voluntary or intentional actions, but is of wider application; it is restricted to the wider field of things the agent knows about *not*

as an observer, so that it includes some involuntary actions; (c) that motives are not mental causes; and (d) that there is an application for 'motive' other than the applications of 'the intention with which a man acts'.

13. Revenge and gratitude are motives; if I kill a man as an act of revenge I may say I do it in order to be revenged, or that revenge is my object; but revenge is not some further thing obtained by killing him, it is rather that killing him is revenge. Asked why I kill him, I reply 'Because he killed my brother'. We might compare this answer, which describes a concrete past event, to the answer describing a concrete future state of affairs which we sometimes get in statements of objectives. It is the same with gratitude, and remorse, and pity for something specific. These motives differ from, say, love or curiosity or despair in just this way: something that *has happened* (or is at present happening) is given as the ground of an action or abstention that is good or bad for the person (it may be oneself, as with remorse) at whom it is aimed. And if we wanted to explain e.g. revenge, we should say it was harming someone because he had done one some harm; we should not need to add to this a description of the feelings prompting the action or of the thought that had gone with it. Whereas saying that someone does something out of, say, friendship cannot be explained in any such way. I will call revenge and gratitude and remorse and pity backward-looking motives, and contrast them with motive-in-general.

Motive-in-general is a very difficult topic which I do not want to discuss at any length. Consider the statement that one motive for my signing a petition was admiration for its promoter, X. Asked 'Why did you sign it?' I might well say 'Well, for one thing, X, who is promoting it, did . . .' and describe what he did in an admiring way. I might add 'Of course, I know that is not a ground for signing it, but I am sure it was one of the things that most influenced me'—which need *not* mean: 'I thought explicitly of this before signing'. I say 'Consider this' really with a view to saying 'let us not consider it here'. It is too complicated.

The account of motive popularised by Professor Ryle does not appear satisfactory. He recommends construing 'he boasted

from vanity' as saying 'he boasted . . . and his doing so satisfies the law-like proposition that whenever he finds a chance of securing the admiration and envy of others, he does whatever he thinks will produce this admiration and envy'¹. This passage is rather curious and roundabout in expression; it seems to say, and I can't understand it unless it implies, that a man could not be said to have boasted from vanity unless he always behaved vainly, or at least very very often did so. But this does not seem to be true.

To give a motive (of the sort I have labelled 'motive-in-general', as opposed to backward-looking motives and intentions) is to say something like 'See the action in this light'. To explain one's own actions by an account indicating a motive is to put them in a certain light. This sort of explanation is often elicited by the question 'Why?' The question whether the light in which one so puts one's action is a true light is a notoriously difficult one.

The motives admiration, curiosity, spite, friendship, fear, love of truth, despair and a host of others are either of this extremely complicated kind or are forward-looking or mixed. I call a motive forward-looking if it is an intention. For example, to say that someone did something for fear of . . . often comes to the same as saying he did so lest . . . or in order that . . . should not happen.

14. Leaving then, the topic of motive-in-general or 'interpretative' motive, let us return to backward-looking motives. Why is it that in revenge and gratitude, pity and remorse, the past event (or present situation) is a reason for acting, not just a mental cause?

Now the most striking thing about these four is the way in which good and evil are involved in them. E.g. if I am grateful to someone, it is because he has done me some good, or at least I think he has, and I cannot show gratitude by something that I intend to harm him. In remorse, I hate some good things for myself; I could not express remorse by getting myself plenty of enjoyments, or for something that I did not find bad. If I do something out of revenge which is in fact advantageous rather

¹ *The Concept of Mind*, p. 89.

than harmful to my enemy, my action, in its description of being advantageous to him, is involuntary.

These facts are the clue to our present problem. If an action has to be thought of by the agent as doing good or harm of some sort, and the thing in the past as good or bad, in order for the thing in the past to be the reason for the action, then this reason shews not a mental cause but a motive. This will come out in the agent's elaborations on his answer to the question 'Why?'

It might seem that this is not the most important point, but that the important point is that a *proposed* action can be questioned, and the answer be a mention of something past. 'I am going to kill him'—'Why?'—'He killed my father'. But if we say this, we show that we are forgetting the course of our enquiry; we do not yet know what a proposed action is; we can so far describe it only as an action predicted by the agent, either without his justifying his prediction at all, or with his mentioning in justification a reason for acting; and the meaning of the expression 'reason for acting' is precisely what we are at present trying to elucidate. Might one not predict mental causes and their effects? Or even their effects after the causes have occurred? E.g. 'This is going to make me angry'. Here it may be worth while to remark that it is a mistake to think one cannot choose whether to act from a motive. Plato saying to a slave 'I should beat you if I were not angry' would be a case. Or a man might have a policy of never making remarks about a certain person because he could not speak about that man unenviously, or unadmirably.

We have now distinguished between a backward-looking motive and a mental cause, and found that, here at any rate, what the agent reports in answer to the question 'Why?' is a reason for acting if in treating it as a reason he conceives it as something good or bad, and his own action as doing good or harm. If you could e.g. show that either the action for which he has revenged himself, or that in which he has revenged himself, was quite harmless or was beneficial, he ceases to offer a reason, except prefaced by 'I thought'. If it is a proposed revenge he either gives it up or changes his reason. No such discovery would affect an assertion of mental causality. Whether in general good and harm play an essential part in the concept of intention it still remains to find out. So far they have only been introduced

as making a clear difference between a backward-looking motive and a mental cause. When the question 'Why?' about a present action is answered by a description of a future state of affairs, this is already distinguished from a mental cause just by being future. Hence there does not so far seem to be any need to say that intention as such is intention of good or of harm.

15. Now, however, let us consider this case:

Why did you do it?

Because he told me to.

Is this a cause or a reason? It appears to depend very much on what the action was or what the circumstances were. And we should often refuse to make any distinction at all between something's being a reason, and its being a cause of the kind in question; for that was explained as what one is after if one asks the agent what led up to and issued in an action. But his being given a reason to act and accepting it might be such a thing. And how would one distinguish between cause and reason in such a case as having hung one's hat on a peg because one's host said 'Hang up your hat on that peg'? Nor, I think, would it be correct to say that this is a reason and not a mental cause because of the understanding of the words that went into accepting the suggestion. Here one would be attempting a contrast between this case and, say, turning round at hearing someone say Boo! But this case would not in fact be decisively on one side or the other; forced to choose between taking the noise as a reason and as a cause, one would probably decide by how sudden one's reaction was. Further, there is no question of understanding a sentence in the following case: 'Why did you waggle your two fore-fingers by your temples?'—'Because *he* was doing it'; but this is not particularly different from hanging one's hat up because one's host said 'Hang your hat up'. Roughly speaking—if one were forced to go on with the distinction—the more the action is described as a mere response, the more inclined one would be to the word 'cause'; while the more it is described as a response to something as *having a significance* that is dwelt on by the agent in his account, or as a response surrounded with thoughts and questions, the more inclined one would be to use the word

'reason'. But in very many cases the distinction would have no point.

This, however, does not mean that it never has a point. The cases on which we first grounded the distinction might be called 'full-blown': that is to say, the case of e.g. revenge on the one hand, and of the thing that made one jump and knock a cup off a table on the other. Roughly speaking, it establishes something as a reason if one argues against it; not as when one says 'Noises should not make you jump like that: hadn't you better see a doctor?' but in such a way as to link it up with motives and intentions: 'You did it because he told you to? But why do what he says?' Answers like 'he has done a lot for me', 'he is my father', 'it would have been the worse for me if I hadn't' give the original answer a place among reasons; 'reasons' here of course conforms to our general explanation. Thus the full-blown cases are the right ones to consider in order to see the distinction between reason and cause. But it is worth noticing that what is so commonly said, that reason and cause are everywhere sharply distinct notions, is not true.

16. It will be useful at this stage to summarize conclusions reached so far. Intentional actions are a sub-class of the events in a man's history which are known to him *not* just because he observes them. In this wider class is included one type of involuntary actions, which is marked off by the fact that mental causality is excluded from it; and mental causality is itself characterized by being known without observation. But intentional actions are not marked off just by being subject to mental causality, since there are involuntary actions from which mental causality is not excluded. Intentional actions, then, are the ones to which the question 'Why?' is given application, in a special sense which is so far explained as follows: the question has not that sense if the answer is evidence or states a cause, including a mental cause; positively, the answer may (a) simply mention past history, (b) give an interpretation of the action, or (c) mention something future. In cases (b) and (c) the answer is already characterised as a reason for acting, i.e. as an answer to the question 'Why?' in the requisite sense; and in case (a) it is an answer to that question if the ideas of good or harm are

involved in its meaning as an answer; or again if further enquiry elicits that it is connected with 'interpretative' motive, or intention *with which*.

17. I can now complete my account of when our question 'Why?' is shewn not to apply. We saw that it was refused application if the agent's answer was 'I was not aware I was doing that' and also if the answer implied 'I *observed* that I was doing that'. There was a third circumstance as well, in which the question would have no application: namely that in which the action is somehow characterised as one in which there is no room for what I called mental causality. This would come out if for example the only way in which a question as to cause was dealt with was to speculate about it, or to give reasons why such and such should be regarded as the cause. E.g. if one said 'What made you jump like that?' when someone had just jerked with the spasm which one sometimes gets as one is dropping off to sleep, he would brush aside the question or say 'It was involuntary—you know, the way one does sometimes jump like that'; now a mark of the rejection of that particular question 'What made you?' is that one says things like 'I don't know if anyone knows the cause' or 'Isn't it something to do with electrical discharges?' and that this is the only sense that one gives to 'cause' here.

Now of course a possible answer to the question 'Why?' is one like 'I just thought I would' or 'It was an impulse' or 'For no particular reason' or 'It was an idle action—I was just doodling'. I do not call an answer of this sort a rejection of the question. The question is not refused application because the answer to it says that there is *no* reason, any more than the question how much money I have in my pocket is refused application by the answer 'None'.

An answer of rather peculiar interest is: 'I don't know why I did it'. This can have a sense in which it does not mean that perhaps there is a causal explanation that one does not know. It goes with 'I found myself doing it', 'I heard myself say . . .', but is appropriate to actions in which some special reason seems to be demanded, and one has none. It suggests surprise at one's own actions; but that is not a sufficient condition for saying it, since one can be a bit surprised without wanting to use such an

expression—if one has uttered a witticism of a sort that is not one's usual style, for example.

'I don't know why I did it' perhaps is rather often said by people caught in trivial crimes, where however it tends to go with 'it was an impulse'. I disregard this use of it, as it has become too much of a set form; and it does not in fact seem strange to be attracted to commit trivial crimes without any need (if there is anything strange, it is only in not being deterred by obvious considerations, not in thinking of doing such a thing). Sometimes one may say: 'Now why did I do that?'—when one has discovered that, e.g. one has just put something in a rather odd place. But 'I don't know why I did it' may be said by someone who does not *discover* that he did it; he is quite aware as he does it; but he comes out with this expression as if to say 'It is the sort of action in which a reason seems requisite'. As if there were a reason, if only he knew it; but of course that is not the case in the relevant sense; even if psychoanalysis persuades him to accept something as his reason, or he finds a reason in a divine or diabolical plan or inspiration, or a causal explanation in his having been previously hypnotised.

I myself have never wished to use these words in this way, but that does not make me suppose them to be senseless. They are a curious intermediary case: the question 'Why?' has and yet has not application; it has application in the sense that it is admitted as an appropriate question; it lacks it in the sense that the answer is that there is no answer. I shall later be discussing the difference between the intentional and the voluntary; and once that distinction is made we shall be able to say: an action of this sort is voluntary, rather than intentional. And we shall see (§25) that there are other more ordinary cases where the question 'Why?' is not *made out* to be inapplicable, and yet is not granted application.

18. Answers like 'No particular reason'; 'I just thought I would', and so on are often quite intelligible; sometimes strange; and sometimes unintelligible. That is to say, if someone hunted out all the green books in his house and spread them out carefully on the roof, and gave one of these answers to the question 'Why?' his words would be unintelligible unless as

joking and mystification. They would be unintelligible, not because one did not know what *they* meant, but because one could not make out what the man meant by saying them here. These different sorts of unintelligibility are worth dwelling on briefly.

Wittgenstein said that when we call something senseless it is not as it were its sense that is senseless, but a form of words is being excluded from the language. E.g. 'Perhaps congenitally blind people have visual images'. But the argument for 'excluding this form of words from the language' is apparently an argument that 'its sense is senseless'. The argument goes something like this: What does it mean?—That they have what I have when I have a visual image. And *what* have I?—Something like *this*.—Here Wittgenstein would go on to argue against private ostensive definition. The next move is to see what is the language-game played with 'having a visual image' or 'seeing in one's mind's eye'. It isn't *just* saying these things—nor can it be explained as saying them with the right reference (this has been shewn by the argument against private ostensive definition). The conclusion is that the language-game with 'seeing' is a necessary part of the language-game with 'seeing in the mind's eye'; or rather, that a language-game can only be identified as that latter one if the former language-game too is played with the words used. The result of the argument, if it is successful, is that we no longer want to say 'Perhaps blind men . . . etc.' Hence Wittgenstein's talk of 'therapies'. The 'exclusion from the language' is done not by legislation but by persuasion. The 'sense that is senseless' is the *type* of sense that our expressions suggest; the suggestion arises from a 'false assimilation of games'.

But our present case is entirely different. If we say 'it does not make sense for this man to say he did this for no particular reason' we are not 'excluding a form of words from the language'; we are saying 'we cannot understand such a man'. (Wittgenstein seems to have moved from an interest in the first sort of 'not making sense' to the second as *Philosophical Investigations* developed.)

Similarly, 'I was not aware that I was doing so' is sometimes intelligible, sometimes strange, and in some cases would be unintelligible.