

University of St Andrews



Full metadata for this thesis is available in
St Andrews Research Repository
at:

<http://research-repository.st-andrews.ac.uk/>

This thesis is protected by original copyright

Abstract of M.Litt. dissertation "Explaining Action"

by Zinaida Lewczuk

Our thesis concerns the explanation of human action. According to the view we call methodological monism, to explain an action is to appeal to some causal law. This means that the explanation of action is similar to the mode of explanation characteristic of the physical sciences.

The methodological monist can adopt two different policies. He can try to translate intentional (mental) language into physical language in the description and explanation of action. In Ch. I we try to show that intentional language is indispensable, and when we try to reduce it to extensional language we have to face the problem of indeterminacy of translation because of the holism of mental terms, and different truth-conditions for intentional propositions. The second possible policy for the methodological monist is to attempt to identify mental states with brain states; in Ch. II we show this to be unsatisfactory since the identification cannot be established on empirical grounds alone, but presupposes some unjustified philosophical assumptions. The proponents of an identity theory need to give a clear account of what is involved in the identity of events and states.

In Ch. III we try to disprove the suggestion that reasons are causes of action. Knowledge of one's own future action and the reasons for it is non-inferential, whereas our knowledge of causes is based on inductive evidence. There is a conceptual connection between reasons and actions, so it cannot be a causal connection. In Ch. IV we suggest that to account for an action we use practical reasoning, i.e. the practical syllogism. We show that its nature is entirely different from theoretical inference, i.e. from a proof syllogism. Our conclusion is that the practical syllogism provides the human sciences with a conceptual framework which can be an alternative to the subsumption-theoretic covering law model characteristic of the natural sciences.

I hereby declare that this dissertation has been composed by myself, that the work of which it is a record has been done by myself, and that it has not been accepted in any previous application for a higher degree.

I was admitted as a candidate for the degree of M.Litt. in October 1977.

I certify that the conditions of the Resolution and regulations have been fulfilled.

J.E.R.Squires (supervisor)

EXPLAINING ACTION

A dissertation for the degree of M.Litt.

in the University of St Andrews

by

ZINAIDA LEWCZUK



TABLE OF CONTENTS

INTRODUCTION	page 1	
Chapter I	THE ATTEMPT TO TRANSLATE MENTAL TERMS INTO PHYSICAL	page 8
Chapter II	THE SEARCH FOR EMPIRICAL CORRELATIONS BETWEEN MENTAL STATES AND BRAIN STATES	page 18
Chapter III	THE THESIS THAT EXPLANATION IN TERMS OF REASONS IS CAUSAL EXPLANATION	page 26
Chapter IV	PRACTICAL REASONING AS INTERPRETIVE SCHEME FOR ACTION	page 52

INTRODUCTION

Our main interest is the explanation of human action. We are told by philosophers that to explain an action is to give the agent's reasons for it. But there is a disagreement about the nature of reasons themselves. According to some philosophers, like Davidson and Pears, reasons are causes of action, so explanation in terms of reasons is a species of causal explanation. Others, like Melden, Kenny and Anscombe, believe that reasons are not causes of action, and therefore that explanation in terms of reasons cannot be causal explanation. To resolve this dispute one way or the other, we need a clear account of the notion of explanation in general. The range of meaning of 'explanation' in science is very wide, and is different in different sciences. What scientists understand by it depends on which model of explanation they accept.

In the physical sciences we have two models of explanation: deductive-nomological and inductive-probabilistic. According to the former, an event is explained when it is covered by a law of nature. Given a law of nature and initial conditions, we can deduce a description of the event to be explained. It is more difficult to specify the explanatory role of the inductive-probabilistic model. On this model to explain an event is to cover it by a probabilistic law. But such a law, as opposed to a causal law, doesn't give an answer to the question 'Why?'. It rather justifies certain expectations and predictions. In the biological and social sciences we meet an

entirely different kind of explanation, the so-called teleological explanation. To explain an event on the teleological model is to appeal to a goal or a result aimed at, for the sake of which the event is said to occur. In other words, a present event is explained by means of a future event, whereas non-teleological models appeal to present or past events, but never to future ones. Various philosophers and scientists take upon themselves the examination and critical analysis of these different kinds of explanation.

Some, like Duhem and Mach, suggest that science (they have in mind physical science) doesn't really explain anything. All it can do is develop calculi intended to forecast phenomena. Duhem characterizes physical theory as follows:

A physical theory is not an explanation. It is a system of mathematical propositions, deduced from a small number of principles which aim to represent, as simply, as completely, and as exactly as possible a set of experimental laws.

(The Aim and Structure of Physical Theory, p.19)

If we regard a physical theory as a hypothetical explanation of material reality, we make it, according to him, dependent on metaphysics. An explanation, then, transcends the methods used by physics. A theory whose aim is to explain reality consists usually of two parts. One is simply a representative part, which classifies the laws. Another part tries to explain the reality underlying the phenomena. The link between these two parts is artificial, for the representative part has developed on its own, and the

explanatory part is a kind of parasite. Duhem is convinced that

it is not to this explanatory part that the theory owes its power and fertility; far from it. Everything good in the theory, by virtue of which it appears as a natural classification and confers on it the power to anticipate experience, is found in the representative part, all of that was discovered by the physicist while he forgot about the search for explanation.

(op. cit., p.32)

When a theory is modified, the representative part enters nearly completely into the new theory, whereas the explanatory part is replaced by another explanation. To show this, Duhem gives the following example. Descartes gave a theory which represents the phenomena of simple refraction. It is based on the constant relation between the sine of the angle of incidence and the sine of the angle of refraction. Descartes also gave an explanation of light effects. Light, according to him, is only an appearance; the reality is the pressure caused by the rapid motions of incandescent bodies within a "subtile matter" penetrating all bodies. There is not, according to Duhem, any connection between the explanation of light phenomena and the representation of the law of refraction. While the latter, even today, forms a major part of elementary optics, the Cartesian explanation of light phenomena has collapsed completely. So, Duhem concludes, we shouldn't judge scientific theories on the basis of their explanatory power, but on the basis of their ability to predict phenomena. This approach can be called a predictivist conception of science.

But according to another view, explanation plays a vital role in the development of science, and cannot be eliminated from it. Although adherents of this view don't want to eliminate explanation from science, nevertheless they are convinced that all the different kinds of explanation can be reduced to one kind, namely, that characteristic of the physical sciences. The main aim of such reduction is to achieve unity of scientific methods. Hempel, who represents this view, says:

The nature of understanding, in the sense in which explanation is meant to give us an understanding of empirical phenomena, is basically the same in all areas of scientific enquiry; and that the deductive and the probabilist modes of nomological explanation accommodate vastly more than just the explanatory arguments of, say, classical mechanics, in particular, they accord well also with the character of explanations that deal with the influence of rational deliberation, of conscious and subconscious motives, and of ideas and ideals on the shaping of historical events. In so doing our schemata exhibit, I think, one important aspect of the methodological unity of all empirical science.

(Explanation in Science and History, p.79)

This emphasis on the unity of scientific methods goes hand in hand with the assumption that the models of explanation characteristic of the physical sciences represent a methodological ideal or standard which measures the degree of development and perfection of all the other sciences, including the humanities. This approach can be called methodological monism. Although there are fundamental differences between the predictivist conception of science and methodological monism, they have one thing in common. Science is regarded according to both as a single phenomenon, which has one goal, namely the forecasting of phenomena, and one method, namely that of the physical sciences.

Both these views can be contrasted with methodological pluralism, represented by such people as Winch, Von Wright, Toulmin and others. The methodological pluralist, even if he agrees with the methodological monist that there is obvious advantage in having a unitary concept of explanation, valid for all branches of science, doesn't believe however that this by itself guarantees that such a unitary concept can be found. The contemporary situation in science ~~has~~ something different. The various sciences are like different games with a great range and variety of rules and purposes. Therefore, says Toulmin, it is:

fruitless to look for a single, all-purpose 'scientific method': The growth and evolution of scientific ideas depends on no one method, and will always call for a broad range of different enquiries. Science as a whole - the activity, its aims, its methods and ideas - evolves by variation and selection.

(Foresight and Understanding, p.17)

Methodological pluralists are convinced that the attempt to reduce teleological explanation to the explanation characteristic of physical sciences is not only unrealistic but also insane. And anyway, we cannot talk about the possibility of such reduction in general, because teleological explanation itself is not just a single kind of explanation but several different patterns.

When we talk about teleological explanation in biology, what we really have in mind is functional explanation. A statement such as "A function of the kidney is the excretion of urine" can be an example of functional explanation. Another case of teleological

explanation is what some philosophers call purposive explanation, which is used to explain goal-directed activity. By the latter we understand a persistence towards the goal under varying conditions. But goal-directed activity understood in such a way covers a very vast range of phenomena: human and animal behaviour on the one hand, and the movements of artificial machines on the other. Although such machines show some observable traits of behaviour, we hesitate to call it behaviour, because we feel that the nature of their movements is completely different from the behaviour of human beings and animals. We are also convinced that there is a difference between the explanations of human and animal behaviour. Human beings can give reasons for their actions, which cannot be ignored in the explanation of their behaviour, whereas in explaining animal behaviour we cannot appeal to reasons. Because of this, some philosophers want to make a contrast between teleological explanation of human behaviour and other forms of teleological explanation, calling the former explanation in terms of reasons. According to Von Wright, such explanation is the only "genuine" form of teleological explanation. Other forms he calls "quasi-teleological" because although they can be couched in teleological terminology, they nevertheless depend for their validity on the truth of nomic connections, unlike "genuine" teleological explanation. Since we are not interested

here in ~~genuine~~ teleological explanation in general,
but in the teleological explanation of human behaviour,
let us concentrate on the possibility of assimilating
the latter to the mode of explanation characteristic
of the physical sciences.

Chapter I - THE ATTEMPT TO TRANSLATE MENTAL TERMS INTO PHYSICAL

In order to eliminate teleology from the explanation of human behaviour, the methodological monist can adopt the following policy. He can try to reduce psychological terms to physical terms in the description and explanation of action. By so doing, he thinks we can achieve unity of the language of science. Some, like Skinner, believe that we achieve even more, namely getting rid of unobservable entities like intentions, motives, etc., propositions about which cannot be verified. In general we get rid of "metaphysics". There are two possible ways of eliminating psychological terms. Firstly we can reduce (some philosophers also talk of translation) psychological terms to physical terms describing bodily movements. Secondly, we can search for states of the brain which correlate in lawlike fashion with beliefs, desires, intentions, etc.. The first task is methodological, concerning the description and explanation of human action. The second seems to be ontological, implying something about what the mental is, and involving the mind-body problem. So let us separate these two tasks and concentrate first on the former.

Any discussion concerning the possibility or impossibility of reducing mental concepts to physical ones should be preceded by an analysis of the very concept of the mental, because it is quite far from clear what philosophers understand by it. Mental terms are supposed to designate phenomena which are non-spatial, private, holistic, intentional, etc..

Not being in space has been regarded, since Descartes' distinction between res cogitans and res extensa, as the main

feature of mental phenomena. But contemporary philosophers are sceptical about this distinction, since there are some physical phenomena, like temperature, which also are not in space. A physical thing has a temperature, and temperature itself is not a thing, it is a property of the thing, like pain is a property of living organisms. So not being in space cannot be a clear criterion of the distinction between the mental and the physical.

It is not very clear what philosophers understand by the holism of the mental. Some suggest that when we explain an action in terms of the agent's beliefs and desires these beliefs and desires do not come individually, but are always tied up together, and usually imply some other phenomena such as memory, rationality, etc.. Suppose John is going to a shop to buy bread. To explain his action we have to pick out not only his desire to buy bread and his belief that he can buy it in this particular shop, but also his desire to buy fresh bread, because stale bread will not do. We also tacitly presuppose that John is rational, that he will not go to a furniture shop to buy bread. The concept of rationality is ineliminable from every explanation of human behaviour. We have to presuppose that, in general, an agent will tend to believe what is true, that he will not tolerate inconsistent beliefs, etc.. Some philosophers argue, however, that holism understood in such a way is not a distinctive feature of mental phenomena. For when we are giving a causal explanation of a physical event we also have to take into account a lot of factors, which constitute a necessary and sufficient condition. If I want, for example, to give a full account why a short-circuit caused a fire in a house I have to mention such

factors as the presence of oxygen, the presence of inflammable material, the absence of a suitably-placed sprinkler, etc.. Although this is true, it is not possible to talk about holism in the latter context. There is a contingent relation between all the factors involved in physical causality, whereas there is a conceptual connection between beliefs, desires, rationality, etc.. It is not possible to ascribe beliefs or desires to a person one by one. A particular belief or desire has its meaning only in connection with other beliefs, preferences, intentions, etc., but taken from the whole context it loses its meaning. This fact generates the so-called indeterminacy of translation, which is a serious problem for the monist who wants to reduce mental descriptions to physical ones.

As far as the criterion of privacy is concerned, there is a lot of confusion. There are different uses of the predicate 'private', but only one is interesting for us, namely, that only an agent can know his mental states or inner experiences, nobody else can. According to Descartes and his followers, private knowledge of inner experience presents no problem; it is as possible as knowledge of public objects. Wittgenstein denies the possibility of such knowledge; he believes that it is impossible for a person to use words which refer to private objects, that is, objects which nobody else could - even in principle - know of (the impossibility here is conceptual). It is impossible because there is no such thing as private knowledge, as there is no such thing as a private language. Knowledge by its very meaning implies that there are criteria for its correctness, like following a rule is a criterion for speaking a language. But such criteria are not available for a speaker of a

private language. Neither memory or anything else can be such a criterion. When my memory is entirely wrong I haven't available any further criterion to correct it. So knowledge of inner experience has to be public, based on observable behaviour, if it is going to be knowledge at all. Some philosophers believe that this means that every statement about mental states can be reduced to statements about observable behaviour or dispositions to behaviour. But despite the fact that knowledge of this kind is public there are still differences between the ways in which an observer and an agent acquire it. An observer's knowledge is inferential, he makes inferences on the basis of observed behaviour, whereas the agent's knowledge is non-inferential, based on direct experience. That is why we don't ask, for instance, how he knows that he is in pain. There is no room for such a question. Does this mean that an agent's knowledge is incorrigible, and that he is the final authority as far as propositions about his mental states are concerned? In some cases, like pain, he is, because an essential feature of pain is the person's feeling it. As far as motives, intentions, etc., are concerned, there is room for error, since in some cases a psychoanalyst can know the agent's motives better than he himself does. By observing an agent's behaviour, the psychoanalyst can make correct inferences and even contradict the agent's own expression of his intentions or motives, when there is an inconsistency between what the agent said and what he is doing or not doing. The inconsistency could be because of the agent's deliberate deceiving of others or his own self-deception. But there are no clear boundary lines between the two. So there is a meaning of 'private',

namely, that the agent's knowledge of his mental states is non-inferential, which can be a distinctive mark of the mental.

According to Brentano, the main mark of the mental, as opposed to the physical, is intentionality. He says:

Every mental phenomenon is characterized by what a scholastic of the Middle Ages called intentional (and also mental) inexistence (Inexistenz) of an object (Gegenstand), and what we would call, although not in entirely unambiguous terms, the reference to a content, a direction upon an object (by which we are not to understand a reality in this case, or an immanent objectivity). Each one includes something as object within itself, although not always in the same way. In presentation something is presented, in judgements something is affirmed or denied, in love (something is) loved, in hate (something) hated, in desire (something) desired, etc..

The Distinction between Mental and Physical Phenomena
pp.50-1.

Some philosophers, like Kathleen V. Wilkes in her book Physicalism (Routledge, 1978), suggest that intentionality cannot be the criterion for sensations. But even if this is right, there is still a large number of mental phenomena such as believing, wishing, hoping, thinking, etc., which are intentional, and it is usually phenomena such as these which we cite in the ordinary explanation of action. So in such explanation we must use intentional language. But we do not use intentional language when we describe and explain physical phenomena. There are the following criteria for identifying an intentional sentence:

- 1) A sentence is intentional if it contains a singular term such that whether that term has reference is independent of the truth-value of the sentence. For example 'John believes in Zeus' - the truth or falsehood of this sentence or its

negation is independent of our judgements about the existence of Zeus. 2) A sentence is intentional if the truth-value of the whole sentence is independent of the truth-value of the sentence it contains. For example, 'John believes that today is Friday', in which 'Today is Friday' can vary without varying the truth-value of the whole sentence. 3) A sentences is intentional if replacing the singular term it contains by a term with a different meaning but the same reference will affect the truth-value of the complete sentence. For example, 'John believes that Aristotle is the author of the Ethics', in which if 'Aristotle' is replaced by 'the disciple of Plato', the truth-value of the whole sentence may change. 4) An intentional sentence may contain a term which does not refer to any specific thing. For example, 'John wants a wife'. So philosophers or psychologists who want to reduce intentional language into an adequate non-intentional language for the description and explanation of human behaviour have to show that it is possible to translate intentional sentences into extensional ones without changing the truth-conditions. Suppose they would like to translate the intentional sentence 'John believe that today is Friday' into an extensional one. The given sentence is intentional because if the word 'Friday' is replaced by a coreferential expression, say 'the first of January', the sentence may change its truth-value. It has been suggested that the sentence will no longer be intentional if it is interpreted as expressing a certain attitude of John towards the sentence 'Today is Friday', rather than towards what is stated by the sentence. Thus we will have as translation 'John believes in (accepts) the sentence "Today is Friday" '. But suppose John is Polish

and doesn't know English. Then although it is true that John believes that today is Friday, it is not true that John accepts the sentence "Today is Friday", since John doesn't understand English and cannot accept any English sentence. Let us then try a different interpretation, like the following: 'John accepts a sentence which is a translation of the English sentence "Today is Friday" into some other language'. But this interpretation will not do the job either. We have to presuppose that John understands the meaning of some sentence, no matter what language he speaks, and understanding itself is an intentional concept. To resolve this dilemma behaviourists adopt the following policy. They try to provide "physical" translations of belief-sentences. Some of them describe beliefs in terms of responses to certain stimuli. On this view to say that a person believes in something is to say that he makes some response in the presence of the object of his belief. But it is already clear that this definition is not good enough because it is possible for a person to believe in something which doesn't exist, so that it is in principle impossible for him to be in the presence of the object of his belief. Suppose the behaviourist modifies his definition and says that a person believes in something if, when in the presence of some stimuli, he makes a response which is appropriate to the object of his belief. In general, a person believes a proposition p if and only if he behaves or is disposed to behave in a way which is appropriate to p. But 'appropriate' is an intentional term itself, which cannot be specified without appealing to a person's beliefs. So this attempt too leads to failure.

Some behaviourists and philosophers like William James try to interpret beliefs in terms of fulfilment and disruption, or satisfaction and surprise. On this view, to say that A believes that X will occur within a certain period means that A is in a bodily state which would be satisfied if and only if X occurs in this period. So if John believes that today is Friday, this means that he is in a state which would be satisfied if and only if today is Friday, and which would be disrupted if today is not Friday. Suppose John looks into his diary and it says that today is Friday. Now, according to the behaviourist, he is in a state which can be called satisfaction. But we have to assume that when John was looking into his diary he believed that it tells the truth, that he believed that the diary can give him that information. We wanted to translate the statement about belief into statements about satisfaction, but to do that we tacitly presupposed other beliefs. We also presupposed that John understands the meaning of the word 'Friday' (given that he speaks English). So this last attempt too is condemned to failure. We do not need to look at other attempts to eliminate intentionality from statements about beliefs in order to see that such an attempt is in principle impossible. In general we have to assume that a person uses some language, and using language implies understanding the meanings of words and sentences; and these references to understanding and meaning imply intentionality. For simplicity, we confined ourselves to the concept of belief, but what we said is true about all intentional concepts. Intentional language cannot, in

principle, be eliminated from the description and explanation of human behaviour.

There are also some other reasons why mental or psychological concepts cannot be eliminated from the explanation of human behaviour. Firstly, an action cannot be identified without reference to intentionality, so its explanation too is impossible without using intentional language. Suppose we want to identify the action of signalling a turn. The methodological monist may say that we can identify it by the bodily movement alone, i.e. extending an arm. Suppose however that there are similar circumstances but a person extends his arm not in order to signal a turn but in order to say hallo to his friend. The bodily movements in both cases are the same but the actions are different. What distinguishes one action from another is the agent's intention. There are also a lot of cases in which we cannot observe bodily movements but still can talk about actions. Suppose I see my friend sitting in a chair, and when I ask him if he is taking a rest his answer is 'No, I am meditating'. It is obvious that in this, and other similar cases, we cannot identify an agent's action on the basis of his bodily movements. So intentions are indispensable for the identification of actions. Since identification of the thing to be explained is essential for explanation, intentions are ineliminable from the explanation of human actions.

Secondly, we cannot leave out in our explanation of actions an agent's own impression of his intentions, desires, etc..

But they are available to him only under psychological descriptions. The agent's account of his desires is important in the explanation of his actions because very often they operate only because he is aware of their existence. This very awareness helps them to operate. The agent's identification of desires, beliefs, etc., as his own enables him to plan the action and control it. In general it is not the agent's actual situation which explains what he does, but his conception of the situation and his picture of himself as an agent. Of course, it may be necessary to refer to the actual situation in explaining his success or failure in translating his intentions into action. But if we eliminate an agent's account of his reasons, and interpret his action only in terms of bodily movements, we cannot talk about the success or failure of his action. Suppose I see someone taking a book from a shelf, but it happened that he took the wrong book. If I describe his action only in terms of bodily movements I will not be able to say that his action was a failure. It follows that if we want to explain his action we must use intentional language.

Chapter II - THE SEARCH FOR EMPIRICAL CORRELATIONS
BETWEEN MENTAL STATES AND BRAIN STATES

We have already mentioned that the methodological monist can adopt another policy, namely, to search for brain states which correlate in lawlike fashion with beliefs, desires, intentions, i.e. mental states. If we find such correlations then it will be possible to describe and explain human actions in purely physical terms. It can also mean that mentalistic psychology can at some stage be reduced to neurophysiology. This has become the main task of some identity theories. There are two kinds of identity theory: some assert identity between types of mental and physical states, so they can be called type-identity theories (Armstrong's and Feigl's theories belong to this group), while others, such as Davidson's anomalous monism, deny the type-identity thesis and assert identity only between particular mental and physical states. Adherents of this token-identity thesis are not methodological monists because they deny the possibility of psychophysical laws and hence of the reduction of psychology to neurophysiology. Since we are interested here in methological monism let us concentrate on the type-identity theory only.

There are different kinds of identification. On the one hand there are analytic identifications, like ' $4 = 2^2$ ', which are true in virtue of definitions and principles of arithmetic or logic. Another kind of analytic identification involves synonymy, like that between bachelor and unmarried man, which is true in virtue of the meaning of the terms. On the other hand we have empirical identity statements

like 'Plato is the author of The Republic', and theoretical identity statements in science like 'Temperature is the mean kinetic energy of molecules'. We are told by adherents of the type-identity theory that identity between the types of mental and brain state is similar to theoretical identification in science. If so, then the type-identity thesis should be confirmed or disconfirmed by empirical facts, and cannot be decided by philosophers alone. But this is not quite so, for the following reasons. When identity is asserted between types of mental and physical state, this can be understood as identity of reference. The two expressions 'mental state of type x' and 'brain state of type y' are said to refer to the same events or states if and only if whenever a person is in a mental state of type x he is also in a brain state of type y. So the factual content of type-identity statements is exhausted by the corresponding correlation statements. When identity is thus understood as identity of reference, it fails to be a significant thesis of identity, because, as Kim pointed out;

Not only the psycho-physical identity statement, but also the corresponding "psycho-physical interaction statement", the corresponding "psycho-physical double-aspect statement", and so on, are all confirmable or refutable by fact. And the very same evidence will confirm all of them or none of them; the very same evidence will refute all of them or none of them.

("On the Psycho-physical Identity Theory",
American Philosophical Quarterly 1966, p.228)

This is a serious problem for advocates of the identity theory. If it has the same factual support as its rivals there is no reason for choosing it rather than any other theory. Some adherents of the identity theory argue, however, that despite this fact there are some philosophical reasons for preferring it. Firstly, they argue, the identity

theory leads to scientific simplicity, because the identification of mental with physical states enables the reduction of mentalistic psychology to neurophysiology. Secondly, it leads to ontological simplicity, to a simpler scheme of entities. H. Feigl writes:

The step from parallelism to the identity view is essentially a matter of philosophical interpretation. The principle of parsimony as it is employed in the sciences contributes only one reason in favour of monism. If isomorphism is admitted, the dualistic (parallelistic) position may be retained, but no good grounds can be adduced for such a duplication of realities or even of "aspects" of reality.

('The "Mental" and the "Physical", p.461, in Minnesota Studies in the Philosophy of Science, Vol.II)

Let us now look more closely at how scientific simplicity can be achieved by the type-identity theory. When one theory is reduced to another the laws of the reduced theory can be deduced from those of the theory to which it is reduced. But the reduced and reducing theories usually have different concepts, so we need so-called "connecting principles", which are statements relating concepts of both theories. If we want to reduce classical thermodynamics to statistical mechanics, for example, a statement such as 'Temperature is the mean kinetic energy of molecules' can serve as a connecting principle. The above identification can be interpreted as follows. Whenever gas has such and such temperature, it has such and such mean kinetic energy, and conversely. The advocates of the type-identity theory suggest that in order to reduce mentalistic psychology to neurophysiology we need similar connecting principles, and they suggest that identity statements between types of mental and brain states can serve. And as we said before, these identity statements are based on nothing more than the

corresponding correlation statements. But such correlation statements are not possible, for the reasons we gave in the previous chapter. We showed that mental concepts are holistic, that they have different truth-conditions from physical concepts, so that it is in principle impossible to reduce one to the other. The identification is also impossible because of the fundamental difference between physical objects or states and mental states. Physical objects have a sort of underlying structure (you can call it real essence if you like). When observable characteristics are not a decisive criterion of identity, we have available a more fundamental criterion, namely their atomic structure. But our psychological terms do not carry the assumption of a real essence, discoverable in principle, which in a case of conflict can help us decide the case one way or another. There is no possible candidate for a non-physical real essence. Suppose that a mental state's real essence could be identified with its causal or functional role. If so, then every type of mental state could be associated with a certain causal role. But this is not possible, since a causal role of a mental state would be specified a priori. Pain would be nothing but the cause of pain behaviour. My intention of going to a cinema will cause my going to the cinema, etc.. This means that new discoveries will not be able to change our preliminary specification of the causal role of mental states. If there is no possibility of regrouping in our classification of mental states their essence cannot be identified with their causal role. The same holds for the identification of mental state's real essence with their functional role.

It is also possible to argue that the question whether a given science is reducible to another cannot be raised without reference to some particular stage of development of the two disciplines in question. It was not possible, for instance, to decide in general whether reduction of thermodynamics to mechanics was possible, before both sciences had reached some stage of development. First of all, the laws of both reduced and reducing sciences have to be confirmed independently of each other before reduction. But the situation is different in the case of psychology and neurophysiology. So far we are not able to establish psychological laws, and the question is whether it is in principle possible to establish any. But this second problem can be discussed in detail later on. What follows here is that the argument from scientific simplicity cannot work.

A different criticism of the type-identity thesis was presented by S.Kripke in his two papers, "Naming and Necessity" in Semantics of Natural Language, edited by Davidson and Harman, and "Identity and Necessity" in Identity and Individualization, edited by Munitz. We are told, says Kripke, that mental states are identical with brain states, for example, that pain is C-fibre stimulation. But 'C-fibre stimulation' and 'pain' are rigid designators, by which he means terms that designate the same object in all possible worlds. And because they are rigid designators, the objects that they designate are, if identical, then necessarily identical. But we can imagine a situation in which a person

has C-fibre stimulation but doesn't feel pain. The possibility that they might come apart is inconsistent with their being necessarily identical. The advocates of the identity theory have to explain this inconsistency away, but according to Kripke they cannot do so.

An opponent who would like to reject Kripke's criticism can argue that the identity of mental states with physical states is accidental, so it couldn't be necessary. Kripke's reply to this argument is as follows. If somebody accepts that every object is necessarily self-identical, then he has to accept that an object which is identical with it is also necessarily identical with it. What made some philosophers believe that some identity-statements between rigid designators are contingent is, according to Kripke, their failure to distinguish between necessity and a prioricity. They wrongly believe that only a priori identity statements can be necessary, whereas a posteriori identity statements must be contingent. Kripke argues that the concepts of necessity and a prioricity are not synonymous. He gives the following example. Suppose I want to know whether such and such a number is prime. This fact can be known a priori, but this doesn't mean that it must be known a priori. A computer, for example, can tell me that the number is prime; I then believe this on the basis of a posteriori evidence. It follows that there are necessary truths which can be known a posteriori. We know a posteriori (it was discovered by astronomers) that the Evening Star is identical with the Morning Star, but if it is true that

they are identical, then they are necessarily identical, because the terms 'the Evening Star' and 'the Morning Star' are rigid designators.

Suppose someone would like to meet Kripke's argument by showing that 'pain' and 'C-fibre stimulation' are not rigid designators. According to Kripke, 'pain' has to be a rigid designator because the object is picked out by its essential property. And it will be absurd to say that that very sensation could have existed without being pain. Suppose 'C-fibre stimulation' is not a rigid designator, then, says Kripke, let us pick out its designation by another property which is really essential to it, for example, a specific type of configuration of molecules. Then this will bring us back to the same problem. So this argument too cannot be sound.

Some philosophers believe that in some cases it is possible for a person not to feel pain, although his pain exists. If so, then pain and the corresponding brain-state always come together, and Kripke's claim that they can come apart is unjustified. To show that pain can exist without the person feeling it, Demos in his article "Self-Deception" (Journal of Philosophy, 1960) gives the following example. Suppose someone with a headache goes to the cinema where his attention is engrossed by an exciting film. Because he is excited by the film he doesn't feel any headache, but as soon as the film is over the headache comes again. According to Demos, we can say that the headache continued in existence all the time, but the person didn't notice it because he was engrossed in the film. Thus pain could exist

without the person feeling it. But this argument is wrong for the same reasons as the previous one. Feeling pain is its essential property, so we cannot say that it could exist without the person feeling it. Kripke is right when he says 'The experience itself has to be this very experience, and I cannot say that it is a contingent property of the pain I now have that it is a pain' ("Identity and Necessity", p.161). Demos has given us no reason to say that the headache existed unfelt, rather than just that the filgoer had a headache before and after, but not during, the film. So Kripke's claim that mental states and brain states are not necessarily identical, and therefore not identical at all, cannot be refuted by this argument.

Kripke's argument sounds quite convincing, but in order for it to be decisive we need a clear account of the identity of events and states; unfortunately, he doesn't provide us with such an account. For this dispute to be settled, both sides need to understand what is involved in the identity of events, what is the criterion for saying that one event or state is identical with another. The proponent of the identity thesis has the stronger obligation to make this point clear.

Chapter III - THE THESIS THAT EXPLANATION IN TERMS OF
REASONS IS CAUSAL EXPLANATION

Other methodological monists, like Pears, try to transform an explanation in terms of reasons into a causal explanation without reducing psychological terms to physical ones. To determine whether such an attempt can be successful, we must first examine the notion of causality itself and make clear what notion of causality is involved here. Aristotle associated causes with unobservable forces or powers which necessarily produce corresponding effects. In the Metaphysics he wrote 'When the agent and patient meet suitably to their powers, the one acts and the other is acted on of necessity' (Book IX, Ch. V). Spinoza believed that the connection between cause and effect is a logical connection of some sort. He said 'Given a determinant cause, the effect follows of necessity, and without its cause, no effect follows' (Ethics Book I, Axiom III). It was Hume who freed Western philosophers from associating a cause either with the obscure notion of unobservable power or with the notion of logical necessity. According to him, the essence of the notion of causation consists in a "regular sequence" or "constant conjunction" between pairs of events in nature. Necessity or power do not exist in nature but are found in human minds, when we are determined to believe in necessity by observing a constant conjunction of two events. Modern philosophers, although they entirely accepted Hume's criticism of the association of causality with logical necessity or power, have not adopted his doctrine as it stands. Instead of talking of constant conjunction of events as the real essence of the notion of causality, they say that every causal statement implies by

its very meaning a general proposition asserting a universal connection between kinds of events; in other words, it implies a law. Modern philosophers also disagree with Hume that a mere constant conjunction of events constitutes the essence of causality. Not all events which follow each other in invariable sequence are causally related, and not every general proposition asserting a sequence between kinds of events is regarded as a causal law. Law-like generalizations can be distinguished from accidental generalizations by the place they occupy within a scientific theory. As R.B.Braithwaite said, a law-statement is one which 'appears as a deduction from higher-level hypotheses which have been established independently of the statement' (Scientific Explanation, p.303). Also, in spite of Hume's criticism of associating causality with necessity, some philosophers believe that causality is some sort of necessary connection after all. But they do not think it is logical necessity, they call it physical necessity and analyze it in terms of counterfactual conditionals. Such counterfactuals justify us in making inferences not merely as to what has happened or will happen, but also as to what would have been the case if some event, which did not in fact occur, had occurred. Counterfactuals themselves are explained on the nomic-inferential model.

Although the above doctrine of causality sounds better than past accounts, it is not without problems. To explain causality it appeals to counterfactuals. But, as Kim showed in his paper "Causes and Counterfactuals" (Journal of Philosophy, 1973), on the one hand counterfactual dependency is too broad to pin down causal dependency, and on the other

hand it is too narrow to capture causal dependency. It is too broad because it captures some sort of logical or analytical dependency, as in the following examples: 'If yesterday had not been Monday, today would not be Tuesday' or 'If my sister had not given birth at t , I would not have become an aunt at t '. Sometimes it also captures an accidental connection like 'If this coin were a dime it would be made of silver'. It is too narrow because it doesn't capture cases of over-determination, such as when two bullets pierce a man's heart simultaneously and they are each independently sufficient for his death. It follows then that it is not enough to say merely that a law-like statement, as opposed to an accidental generalization, is a statement which implies a counterfactual. It has to imply proper counterfactuals. But in order to specify which counterfactuals are proper we have to appeal to laws, and say that they are those which are supported by laws. But such an explanation cannot be satisfactory because it is circular. If we are not able to give a satisfactory account of the notion of causation maybe we should give up using the notion. Some philosophers think that instead of talking about causal laws we can talk about probabilistic laws. But Inductive-Probabilistic models of explanation are far from being unproblematic, so they cannot be a simple alternative to causal explanation. Being aware of these problems involved in causal explanation, we can now concentrate on a more concrete problem, namely everyday life explanation in terms of causes.

In ordinary life we are not so much interested, as in the physical sciences, in types of occurrence which usually

or normally occur, on the contrary, everyday causal questions are inspired by the wish to explain the particular events whose occurrence is puzzling because it is a departure from the normally expected course of events. In ordinary life we make singular causal statements. The question is what meaning of 'cause' is presupposed here, and on what basis are we justified in making such statements. If the meaning of 'cause' lies, as Hume thought, in the constant conjunction of events or in a kind of regularity as modern thinkers suppose, then it follows that we cannot observe regularity in an individual case. In ordinary life we also seem to make inferences about effects from causes without any appeal to necessity. How else, if by appealing neither to universality nor to necessity, can we grasp the notion of cause in everyday life?

According to Anscombe, knowledge of causes is possible without any satisfactory analysis of what is involved in causation. In her paper "Causality and Determination" (Inaugural Lecture, Cambridge University Press, 1971), she writes:

How does one show someone that he has the concept cause? We may wish to say: only by having such a word in his vocabulary. If so, then the manifest possession of the concept presupposes the mastery of much else in language. I mean: the word 'cause' can be added to a language in which are already represented many causal concepts. A small selection: scrape, push, wet, carry, eat, burn, knoek over, keep off, squash, make (e.g. noises, paper boats), hurt. But if we care to imagine languages in which no special causal concepts are represented, then no description of the use of a word in such languages will be able to present it as meaning cause. Nor will it even contain words for natural kinds of stuff, nor yet words equivalent to 'body', 'wind', or 'fire'. For learning to use special causal verbs is part and parcel of learning to apply the concepts answering to these, and many other, substantives.

Even if Anscombe is right that the ordinary man can make and understand causal statements without being able to specify the meaning of the term 'cause', it is the philosophers' business to make analytically explicit the meaning of the term as used in ordinary language.

According to Davidson, the ordinary meaning of 'cause' is not fundamentally different from the scientific one. When in ordinary life we say that a caused b we believe that there exists some causal law, even if we don't know what it is, in virtue of which a is a necessary and sufficient condition for the existence of b.

Let us quote his own words:

And very often, I think, our justification for accepting a singular causal statement is that we have reason to believe an appropriate causal law exists, though we do not know what it is. Generalizations like 'If you strike a well-made match hard enough against a properly prepared surface, ~~then~~, other conditions being favourable, it will light' owe their importance not to the fact that we can hope eventually to render them untendentious and exceptionless, but rather to the fact that they summarize much of our evidence for believing that full-fledged causal laws exist covering events we wish to explain.

(*"Causal Relations"*, Journal of Philosophy, 1967)

To say that a singular causal statement implies that there is some corresponding law, even if we cannot know it, is not an obvious thing to say; so in order to make us believe in his thesis Davidson should give us some more reasons.

Mackie, who like Davidson thinks that singular causal statements can be analyzed in terms of necessity and law-like regularity , presents a better elaborated and justified thesis, so let us look at his theory. He goes through the following stages in his analysis.

A cause of an event is said to be an INUS condition of it, and the notion of INUS condition is interpreted in terms of necessary and sufficient conditions, which are in turn analyzed in terms of counterfactuals, which follow from a corresponding universal law. Mackie defines a cause of an event as follows: 'the so-called cause is, and is known to be, an insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result' ("Causes and Conditions", American Philosophical Quarterly, 1965). For example, in saying that a short-circuit was the cause of the fire in the house we mean that the short-circuit was an insufficient but necessary part of a condition such as the presence of inflammable material, the absence of a suitably-placed sprinkler, etc., which itself was sufficient, although unnecessary for the occurrence of a fire. But since the fire might have been caused not by the short-circuit, but by the throwing of a lighted cigarette or in some other way, we have to add that no other sufficient condition of the house's catching fire was present on this occasion. Let A stand for the short-circuit, B for the presence of inflammable material, and C for the absence of a suitably placed sprinkler. The conjunction ABC represents a sufficient condition of the fire which contains no redundant factor and is called by Mackie a minimal sufficient condition for the fire. Let DEF, GHI, etc., be all the other minimal sufficient conditions. Then, says Mackie: 'the formula 'ABC or DEF or GHI or ...' represents a necessary and sufficient condition for the fire, each of its disjuncts

such as ABC, represents a minimal sufficient condition, and each conjunct in each minimal sufficient condition such as A, represents an INUS condition' (loc. cit.).

Now it is clear what Mackie means by a necessary condition.

It is the disjunction of all the sufficient conditions, or sometimes the only sufficient condition (when the effect can be produced by only one cause). 'Necessary condition' and 'sufficient condition' are in turn explained on the model of counterfactual statements. The statement 'A short-circuit here was a necessary condition of a fire in this house' is related to the counterfactual conditional 'If a short-circuit had not occurred here, this house would not have caught fire', and the statement 'A short-circuit here was a sufficient condition of a fire in this house' is related to the factual conditional 'Since a short-circuit occurred here, this house caught fire'. Thus to relate singular statements about necessity and sufficiency to non-material conditionals is to refer indirectly to certain universal propositions. 'Thus, if we said that a short-circuit here was a necessary condition for a fire in this house, we should be saying that there are true universal propositions from which, together with true statements about the characteristics of this house, and together with the supposition that a short-circuit did not occur here, it would follow that the house did not catch fire' (loc. cit.). And, similarly, to say that a short-circuit here was a sufficient condition for a fire in this house is to say that there are true universal propositions from which, together with true statements about the characteristics of this house, and

the statement that a short-circuit occurred, it follows that a fire occurred. Although Mackie's theory needs some refinements, as Kim showed in his paper "Causes and Events: Mackie on Causation" (Journal of Philosophy, 1971), most philosophers agree that it constitutes a notable contribution to the analysis of causation in general, and of singular causal statements in particular.

Now let us concentrate on the problems concerning the possible application of Mackie's regularity theory of causation to the explanation of human actions. Mackie's theory says that every singular causal statement implies, by its very meaning, a general proposition asserting a universal connection between kinds of events. Given that, and the assumption that explanation in terms of reasons is causal explanation (Davidson's view), it follows that statements citing reasons are intelligible only by reference to some corresponding lawlike generalization. Some philosophers find this conclusion unacceptable. According to Hart and Honore, 'the statement that a person acted for a given reason does not require for its defence generalizations asserting connexions between types of events. ... It is no part of the meaning of such a statement that if the same circumstances recurred he would do the same' (Causation in the Law, Oxford, 1959, p.21). When we make assertions about reasons we appeal to rationality or intelligibility or both but not to some general law. Rationality, given the reasons and circumstances, tells us what is or was the appropriate thing to do. For example if I want a book, and the only way to get it is to borrow a copy

from my friend, then the rational thing for me to do is to ask my friend to lend it to me. Rationality in this context consists in a successful choice of means to end. It implies a teleological explanation which has logical peculiarities of its own (which we do not want to analyze here) and which is entirely different from causal explanation.

However, advocates of the opposite view can argue that even in the above cases, when we try to explain an action in terms of reasons, we have to appeal to some sort of generalization after all. They are right, but the sort of generalization required is altogether different from that involved in causal accounts. It is true, for example, that given some reasons and circumstances on one occasion and similar reasons and circumstances on another, we expect that a person will act in a similar way. But this generalization is not based on a causal law which says that given these reasons and this situation a person will act in such and such a way; it is more like a recipe which says that, given these circumstances and this end, a rational person would (or should) use such and such means to his end. Let us quote Hart and Honore again:

The general knowledge used here is knowledge of the familiar 'way' to produce, by manipulating things, certain types of change which do not normally occur without our intervention. If formulated, they are broadly framed generalizations, more like recipes in which we assert that doing one thing will 'under normal conditions' produce another, than statements of 'invariable sequence between a complex set of specified conditions and an event of the given kind'. (Causation in the Law, p.29)

If the explanation of actions in terms of reasons were based on lawlike generalizations, then we should be able to bring it under the pattern of explanation characteristic of physical sciences, i.e. the nomological-deductive model of explanation. This consists of two parts, the so-called explanans and explanandum. The former contains two sets of premises: a set of singular statements describing relevant initial conditions, and a set of general laws. The explanandum describes the phenomenon to be explained. This can be represented by the following schema:

$C_1, C_2, \dots C_n$ statements of initial conditions

$L_1, L_2, \dots L_k$ general laws

E description of the empirical phenomenon to be explained

For this schema to be an adequate explanation, it must satisfy various conditions, one of which is that the explanandum must be a logical consequence of the explanans.

But when we try to explain actions on this nomological-deductive model the following difficulty will arise. Very often a fuller description of an action can be at the same time its explanation, whereas the deductive-nomological model requires that the description of the explanandum should be separated from its explanans. To see this more clearly, let us look at an example given by Anscombe in her book Intention (Blackwell, 1957). We see a person moving his arm up and down; when we ask what he is doing, a fuller description of his action 'He is operating the pump' may be an answer to our question. To the question why he is

operating the pump the answer may be 'He is replenishing the water supply'. By giving the further description 'He is poisoning the inhabitants' we can explain why he is replenishing the water supply. We have here one action with four descriptions, and each is related to the next as a description of means to an end. The last description is treated as an end which explains the agent's intention. It is clear that we cannot pick out one of these descriptions as an explanandum, and another as part of its explanans (an initial condition). For they describe the same event, whereas explanans and explanandum are supposed to describe different events. So human actions cannot be explained on the nomological-deductive model.

Further proof that explanation in terms of reasons cannot be causal explanation can be derived from the fact that an agent's own knowledge of his action and his reasons for it is direct or non-inferential, in no way dependent on a lawlike generalization. Firstly, let us concentrate on the thesis that the agent's knowledge of his action, given his beliefs and desires, is non-inferential. We have to assume of course that he knows that it is possible for him to perform the action, that circumstances will not prevent him from acting. This knowledge is inferential. But given it, and his reasons, he knows in a non-inferential way what action he intends to do. Some philosophers suggest that this knowledge of one's future action is based on the observation of one's past behaviour. One observes that in the past when he had such and such reasons one behaved in such and such a way, and on the present occasion, because

one has similar reasons one infers that one will behave similarly. This would make the agent's knowledge of what he is going to do inferential after all. But this view cannot be right. If the agent's knowledge of his action were based on the observation of previous cases, it would follow that in a situation where he has some reasons for the first time he wouldn't know what he is going to do, which is not true (if he has decided what he is going to do). In general, philosophers who believe that this kind of knowledge is inferential misunderstand the very nature of the agent's knowledge of his future action. They assume that it is similar to the prediction of a future event. And certainly, such a prediction has to be based on an inductive inference. Whereas the agent's declaration that he will perform his action, will go to the cinema (for example) should be taken as a sign that he has made up his mind or that he gave something like an order to himself, but it shouldn't be taken as a prediction.

Our second thesis is that the agent's knowledge of his reasons, i.e. his beliefs and desires, is non-inferential too. Philosophers who disagree with our thesis, like Pears, suggest that the degree of feeling concomitant with a particular desire can be evidence for its existence (in "Desires as Causes of Actions", in Questions in the Philosophy of Mind, Duckworth, 1975). But this cannot be true either. It is not clear at all how a person can know that this particular degree of feeling is sufficient for his having a desire. If such knowledge were possible it would be as mysterious as non-inferential knowledge itself. Pears also

assumes that a person can identify his degree of feeling in a non-inferential way, but then it is not clear why he cannot identify his desire in the same way. In general, if non-inferential knowledge were not possible, then an alternative will either imply an infinite regress or a possibility of giving grounds which are themselves unknowable.

A more powerful criticism of non-inferential knowledge was given by Wittgenstein. To show such knowledge impossible he gives the following arguments. He suggests that first-person statements about inner experience are not propositions at all. They are not used to assert something about the world but are manifestations of its being so, they are extensions of natural expressions. For example, 'I am in pain' is like moaning or crying. If so, then to say that I know I am in pain cannot be a proposition either. It either means the same as 'I have pain' or is senseless. 'It can't be said of me at all (except perhaps as a joke) that I know I am in pain. What is it supposed to mean - except perhaps that I am in pain' (Philosophical Investigations, 246). Wittgenstein seems to suggest that not only first-person statements about one's sensations, but also statements about one's intentions, desires, etc., should be assimilated to natural expressions. He says 'What is the natural expression of an intention? - Look at a cat when it stalks a bird, or a beast when it wants to escape' (P.I., 647). He also writes:

Am I to say that any one who has an intention has an experience of tending towards something? That there are particular experiences of 'tending'? - Remember this case: if one urgently wants to make some remark, some objection, in a discussion, it often happens that one opens one's mouth, draws a breath and holds it; if one then decides to let the objection go, one lets the breath out. The experience of this process is evidently

the experience of veering towards saying something.
Anyone who observes me will know that I wanted to
say something and then thought better of it.

(P.I., 591)

But this expressive thesis is true neither in the case of sensations nor in that of intentions and desires. When I make statements about my experience in the past or future tense, like 'I was in pain' or 'I had the intention', they cannot be extensions of natural expressions, so present-tense statements of this kind can hardly be either. 'I am in pain' can be used in a molecular sentence like 'I am in pain, therefore I must be ill', and in this case it has a truth-value so it cannot be a kind of natural expression. I can also talk about my pain as being sharp or dull, and if 'I am in pain' is not an assertion, how could such descriptions be used? So criticism of non-inferential knowledge cannot be based on the expressive thesis, because the latter is wrong.

Wittgenstein also gives different reasons why present-tense psychological statements in the first person should be excluded from the domain of cognitive statements. He suggests that they haven't truth-values because we haven't any criteria for determining such truth-values. So they cannot be an object of possible knowledge, because knowledge requires possession of a true description of a state of affairs which is independent of what it describes. Although this argument sounds quite persuasive it cannot be right either. Sometimes I can say 'I intend to go to the concert' and mean it to be a lie, so then I know that it is false. But if

it can be false, it can also be true. If the thesis that first-person statements about experience do not have truth-values were right, such a lie wouldn't be possible. So the non-cognitive thesis cannot rest on this ground.

Sometimes Wittgenstein suggests that knowledge is possible only where doubt makes sense. But it is senseless to doubt such statements as 'I am in pain' or 'I want to buy a book', so they should be excluded from the domain of knowledge. The premise of this argument is right, for there isn't room for doubt in the above cases, but the conclusion doesn't follow. It is equally senseless to doubt that I exist, or that $2 + 2 = 4$, but it is still possible to talk about knowledge. So Wittgenstein has not shown that non-inferential knowledge is impossible.

Such knowledge is possible because terms like 'pain', 'intention', 'desire' belong to a public language. I learn to apply them to other people on the basis of their behaviour, and I also learn that other people ascribe them to me on the basis of my behaviour. And when I become a master of the language, I ascribe them to myself. Hacker, in his book Insight and Illusion (O.U.P., 1972), writes:

'I am in pain' has, after all, a structure. As such it is complex or articulated. My use of this sentence is only intelligible in so far as I know what 'pain' means, i.e. know how to apply the predicate on the basis of those criteria which constitute its meaning. For this, as we have seen, is a precondition for my ascribing it to myself without criteria. (p.266)

The explanation of actions cannot be causal explanation for another reason. The description of a cause cannot stipulate that its alleged effect should actually occur, but when we cite the reasons for an action we presuppose that the action will occur. This argument, called by some philosophers "the logical connection argument", seems to be the core in the dispute between philosophers who are convinced that explanation in terms of reasons is not causal explanation and their opponents. Despite the frequency with which it has been used it remains pretty obscure. It is not clear what kind of logical or conceptual connection there is between actions and reasons. Melden writes, in his book Free Action (Routledge 1961), p.141:

It is logically impossible that one should want but never do, or that one should never do what one wants to do. For where there is wanting there is an agent who does; where there is doing there is an agent who can be and has been attentive to what he is doing and who acts as he does for a reason; where there is doing for a reason there it is false that one does what one does not want to do.

Melden seems to suggest here that desires or intentions cannot exist without actions because without actions it wouldn't be possible to ascribe such states of mind to human beings. We cannot imagine such a world in which human beings have desires, motives, etc., but never act on them. Nevertheless it is true that on a given occasion a person may want something yet do nothing in order to get it. We can make an analogy with orders here. As Wittgenstein rightly pointed out, although on some occasions it is possible to disobey an order, in general the very notion of an order presupposes that it should be obeyed. And similarly the notion of a reason for action presupposes that of action. This general formul-

ation of the argument sheds some light on the nature of the connection between reasons and actions. But in order to make it clear we have to concentrate on a particular case.

Suppose a person wants to go for a walk, and his desire is followed by the action, i.e. going for a walk. The agent's desire is described in such a way that it mentions the action which it is supposed to cause. From this fact, some philosophers believe, it follows that the desire cannot be the cause of the action. But their conclusion doesn't seem to be justifiable. Firstly, in a lot of cases it is perfectly possible to describe a desire in such a way that it doesn't mention the action by which it is followed. Suppose a person wants a book, and his desire is followed by his going to a library. In this case the description of the desire doesn't mention the action. So if we accept the above criterion it will follow only that some desires are not causes of actions, namely, those which mention an action in their description. But for other desires which do not thus mention actions it is possible for them to be the causes of actions. Secondly, this criterion is too wide. It will exclude some genuinely causal connections, like that between exposure to sun and sunburn.

Suppose we give a stronger criterion, namely that a desire cannot be regarded as the cause of an action if the only way to specify it is by mentioning the action, directly or indirectly. But this criterion doesn't seem to be good enough either. It is still possible to find some causal

statements which violate it. We are told by psychologists that fear of stammering causes some people to stammer. In this case the only way to specify the cause is by mentioning its effect, but this fact still doesn't stop us regarding it as a cause.

In order to prove that desires are not causes of actions we need a better criterion. We can formulate it as follows. A desire cannot be a cause when an specification of the cause stipulates that its supposed effect should actually occur. Then the statement will be analytically true, so it cannot be a causal statement. To understand this criterion better let us give an example. Suppose a person wants to visit his friend on 1-3-79 in the afternoon, that he doesn't change his mind, and that nothing happens to prevent him making the visit. Then it is logically impossible for him not to visit his friend. If he didn't do it we will deny that he really wanted to. It is true in virtue of how the concept of desire is used in our language. Does this connection exclude a causal connection? Pears suggests that sometimes a cause can be specified in such a way that it stipulates that its effect will occur. In his paper "Desires as the Causes of Actions" (loc. cit.) he gives such an example: 'A person's death was caused by a fatal dose of barbiturate' (p.88). But this specification can be removed because there is another specification, namely a particular dose, say forty grains, which doesn't stipulate any more that the effect should actually occur. But when we remove the specification which stipulates that a practical desire should cause the relevant action, we have no other line onto it. Here we have the crux

of the logical connection argument. It says that the criterion for the existence of a desire is the fact that it actually brings about the relevant action, and this is the only criterion. But if so, a desire cannot be the cause of the action, since there should be a contingent relation between cause and effect.

Suppose an opponent denies the premisses of our argument and suggests that a desire cannot be an analytically necessary condition for the relevant action, since there are well-known cases when people want to do something but don't actually do it, or when people want something but don't do anything to get it. This means that there must be a causal explanation of the non-fulfilment of the agent's desire. But we are not denying that; we agree that in a lot of cases there are interfering factors which can prevent the desire from being realized. In these cases there will be a causal explanation of the non-fulfilment. But if they are absent it will be a *priori* necessary that the desire will be followed by the relevant action.

Pears gives a stronger objection against the logical connection argument. He suggests that the operation of a practical desire can be the subject of empirical confirmation or disconfirmation. But this could not be possible if the thesis that there is a logical connection between desire and the relevant action were true. In order to prove that a practical desire can become the subject of empirical test Pears tries to specify the conditions which should be satisfied on any testing occasion. These conditions are psych-

ological and physical. The first condition requires that the agent's desire to do something, all things considered, must not have vanished in the interim. The second requires that the agent must be aware that he is able to perform the action. The physical condition must guarantee that there are no external obstacles to his performance of it. If all these conditions are satisfied then the agent's sincere statement that he will do something should, according to Pears, entail his doing it. So far we have been considering a very simple case, where the agent wanted to do A all things considered, and this was his only desire. But very often we have cases where the desire is compound. The agent wants to do A in order to achieve B and C, for example. There are also cases where a component desire is an aversion. The agent does A in order to achieve B although he doesn't want to bring about C. Because the testing situation is more complicated here we have to specify some further conditions, namely that the component desires engaged on the original occasion must not have decreased in number or in degree in the interim; the agent must know that the situation at the time of the test possesses the same features as on the original occasion; and no aversion must have developed in the agent in the interim. It is clear that the list of requirements to be met by any testing situation cannot be completed; it will always be an open-ended disjunction. If so, then we can always suggest that there is some further factor which should be taken into account in the test, and thus falsifiability is not possible. But Pears has a reply to this point. He suggests that the above fact creates

difficulties not only for confirmation or disconfirmation of practical desires but also for singular causal statements:

It is undeniable that many singular causal statements about physical events, like the one about the billiard balls, have general implicants which, at some degree of specificity and involvement in the details of the particular case, become uncompletable.

("Rational Explanation of Actions and Psychological Determinism", in Essays on Freedom of Action, edited by T. Honderich, Routledge, 1973, p.115)

Pears is right in suggesting that the list of interfering factors can be uncompletable in both cases. But he is wrong in assuming that this fact creates the same problem for testing the operation of a desire and the operation of a cause. Suppose we make a hypothesis about some physical event, say that this match will light when scratched, but that when we scratch it, it doesn't light. We try to find a factor or factors which prevented it from lighting. These could be an absence of oxygen, or the match might have been wet or badly made, or we didn't scratch it hard enough, etc.. This list of preventing factors could be open-ended, but this doesn't stop us from believing that if the conditions were right the match would have lighted. We believe this on the basis of the causal law which says that scratched matches will light if the right conditions are satisfied. Suppose we make a similar hypothesis about a future action, say that the agent will go to a library this afternoon because he says he wants to, but he doesn't actually go. We try to find what prevented him from going - he might have forgotten about his desire, he could have changed his mind, perhaps somebody told him that the library is closed, etc.. Maybe we cannot find which factor prevented him from going. Suppose we ask him why he didn't go, but he is not able

to give us a satisfactory account, yet still claims that he wanted to go. In this case we will deny that he really wanted to. But we cannot deny that the match was scratched. The difference between the two cases lies in the fact that in the first we were able to specify the cause and the effect independently of each other. And although the effect didn't happen and we were not able to explain why, we still don't deny the operation of the cause, i.e. the scratching of the match. But in the second case, the effect was the criterion for the existence of the desire, so when the action didn't happen we denied the existence of the desire. This example shows that the connection between a desire and the relevant action cannot be contingent, and if so it cannot be subject to empirical test. So Pears is trying to do the logically impossible, namely to subject statements which are true in virtue of their meaning to empirical test.

Pears also adopted another policy to prove that the logical connection argument is wrong. He tries to show that an action is not the only criterion for the existence of the relevant desire. Firstly, he proposes the degree of feeling concomitant with a desire as a criterion for its existence. But we have already showed, in our analysis of non-inferential knowledge of one's own desires, intentions, etc., that such identification is not possible. Secondly, Pears suggests that a brain state could be an independent criterion for the existence of a desire. To prove this he cannot appeal to the type-identity thesis, because it doesn't hold; the only thing he

can do is appeal to the token-identity thesis, i.e. to Davidson's anomalous monism. But it is not certain that this thesis is true either. Some philosophers, such as McGinn, suggest that the arguments given against the type-identity thesis are not valid against the token-identity view. He writes:

I agree that this point is powerful against the type-type theorist who accepts Kripke's conditions on an adequate reply to the Cartesian. But I want to insist, against Kripke, that his favoured style of explanation of the impression of contingency is available to the token-token theorist.

(*"Anomalous Monism and Kripke's Cartesian Intuitions"*, Analysis Vol.37, 1976-7, p.79)

To recall, Kripke argues that if types of mental state are identical with corresponding types of physical state they should be necessarily identical, since this is an identity between rigid designators. But we can imagine that they can come apart, which is inconsistent with their being necessarily identical. Why is McGinn convinced that a similar argument cannot be given against the token-identity thesis? Let 'A' be the name of a mental token, e.g. Jones' feeling pain at noon on 1-10-78, and let 'B' be the name of a physical token, e.g. Jones' brain state at noon on 1-10-78. Kripke is convinced 'that it is at least logically possible that B should have existed (Jones' brain could have been in exactly that state at the time in question) without Jones feeling any pain at all, and thus without the presence of A' (*Naming and Necessity* p.335). We can also imagine, he claims, the converse situation, i.e. Jones' feeling pain then without the corresponding brain state. But this possibility that A and B can come apart is inconsistent with their being necessarily identical. McGinn

argues that Kripke is wrong in suggesting that A and B can come apart. If they are necessarily identical it is logically impossible that they can come apart. To understand McGinn's point better let us give an example which is taken from Kripke's own paper "Identity and Necessity", although he doesn't seem to be aware of the possible analogy between it and the token-identity of mental states with brain states. Suppose there is a table in front of me which is made of wood. According to some philosophers, in another possible world this very table could have been made of ice, but Kripke argues that they are wrong. It is an essential property of this table that it is made of wood, so in any other possible world it has to be made of wood too. When we are wondering whether "it" could have been made of ice, we, according to Kripke, talking about a different table. McGinn writes:

In this respect, token mental states are like particular tables, they can be (and be essentially) of a type such that other tokens of that type fail to have properties which they, qua ^{tokens}_A, necessarily have. (loc. cit., p.80)

So Kripke's claim that A and B can come apart is not justified, if McGinn is right. But to resolve this problem one way or the other we need an account of the notions of essential property and of identity across possible worlds (these are equivalent notions). But neither Kripke nor McGinn give a satisfactory account of them, so our problem will have to be resolved on different grounds.

First of all, we will have to make clear what is the nature of the identification of particular mental states with corresponding brain states. Is it an empirical identification, like that of temperature with mean kinetic energy of molecules?

The answer must be no, because temperature and mean kinetic energy can be measured empirically and established independently of each other. Whereas mental predicates do not denote natural kinds, so they cannot be specified empirically. Wittgenstein said:

Misleading parallel: psychology treats of processes in the mental sphere, as does physics in the physical. Seeing, hearing, thinking, feeling, willing, are not the subject of psychology in the same sense as that in which movements of bodies, the phenomena of electricity, etc. are the subject of physics. You can see this from the fact that the physicist sees, hears, thinks, about, and informs us of these phenomena, and the psychologist observes the external reactions (the behaviour) of the subject. (Philosophical Investigations, 571)

If we cannot observe or measure mental states what is our justification for saying that they are identical with brain states? To justify this the adherents of identity theories give us a metaphysical thesis which says that every mental state is identical with a corresponding brain state, and suggest that it can be justified on purely philosophical grounds. But philosophical arguments can also be given in favour of other theories, like double-aspect or parallelist views, which don't identify mental states with physical states. Suppose the advocates of identity theories argue that the identification can be established indirectly, on the basis of behaviour. But if so, then we cannot remove behaviour from the specification of desires, and Pears' attempt to show that the logical connection between desire and action can be removed by using a brain state as an independent criterion for the existence of a practical desire must fail. Pears might then argue that behaviour can be described and explained in purely physical terms, i.e. in terms of mere bodily movements. But we cannot do that without changing

the subject; our description and explanation would no longer be an explanation and description of intentional action. In the previous chapter we argued in detail why the psychological description of intention action cannot be reduced to physical description. So the token-identity thesis cannot serve Pears' purpose, firstly, because it cannot help him remove the logical connection between desire and action, and secondly, because the theory itself is not well-established.

Chapter IV - PRACTICAL REASONING AS INTERPRETIVE
SCHEME FOR ACTION

We have tried to show that various attempts to assimilate explanation in terms of reasons to the explanation characteristic of the physical sciences lead to failure. To explain human actions we use intentional language, which in principle cannot be reduced to the extensional language used in the natural sciences. We also showed that explanation in terms of reasons doesn't depend for its validity upon the truth of nomic connections (causal laws), so the models we use to account for physical phenomena cannot be applied to human behaviour. To account for an action we use practical reasoning, i.e. a practical syllogism. The practical syllogism is regarded by some contemporary philosophers, like Anscombe, Von Wright, Kenny and others, as the best way to explain intentional actions. The conception of the practical syllogism was introduced for the first time by Aristotle and, according to Anscombe, it was one of his best discoveries, although she thinks that it "has been lost to modern philosophy through misinterpretation" (Intention, p.58). Von Wright too is convinced that practical reasoning is of great importance in the explanation and understanding of action:

It is a tenet of the present work that the practical syllogism provides the sciences of man with something missing from their methodology: an explanation model in its own right which is a definite alternative to the subsumption-theoretic covering law model. Broadly speaking, what the subsumption-theoretic covering law model is to causal explanation and explanation in the natural sciences, the practical syllogism is to teleological explanation and explanation in history and the social sciences.

(Explanation and Understanding,
Routledge, 1971, p.27)

What is the nature of practical reasoning? We can say, in general, that it is the kind of reasoning involved in rational deliberation leading up to and providing grounds for an action. In practical reasoning, as in theoretical, we pass from premisses such as statements of the agent's intentions or desires, and statements about the available means, to a conclusion about what he will or must do, given these desires and circumstances.

According to Aristotle:

We deliberate not about ends but about means. For a doctor does not deliberate whether he shall heal, nor an orator whether he shall persuade, nor a statesman whether he shall produce law and order, nor does anyone else deliberate about his end. They assume the end and consider how and by what means it is to be attained; and if it seems to be produced by several means they consider by which it is most easily and best produced, while if it is achieved by one only they consider how it will be achieved by this and by what means this will be achieved, till they come to the first cause, which in the order of discovery is last. ... And if we come on an impossibility, we give up the search, e.g. if we need money and this cannot be got; but if a thing appears possible we try to do it.

(Nicomachean Ethics III, 1112b18ff)

But what Aristotle said about ends doesn't seem to be always true. In some cases a doctor may deliberate whether or not to heal a very senile patient with pneumonia, for example, although he has means available. Sometimes we wonder whether our end is worth our energy and time, and if we come to the conclusion that it isn't we give it up rather than adopt the available means for its realization. Very clear examples of deliberation about ends occur in cases of ^{con}insistent desires. Because we know we cannot satisfy both, we deliberate about which one is more important for us. Aristotle thought that a proposition constituting a universal premise expresses some general human

value, as in the following example:

Dry food suits any human.
 Such-and-such food is dry.
 I am human.
 This is a bit of such-and-such food.
 So, this food suits me.

Contemporary writers suggest that practical reasoning is not ethical argument, so we should not talk of goodness but rather of desirability (Anscombe), satisfaction (Ross) or satisfactoriness (Kenny) as the value to be transmitted by practical reasoning. And these values are not something which can be established objectively, but are relative to the wants of a given individual, since what is desirable or satisfactory for some people might not be so for others. It can also happen that a person can recognise something as desirable in general, like healthy food for example, and yet not accept it as a reason for his action in a particular situation; if he is very hungry, he may eat food which is not healthy. Contemporary philosophers also disagree with Aristotle's assumption that the major premise can be expressed in a universal statement. They suggest that a universal proposition cannot become a starting point for reasoning about what to do in a particular case. To show this, Anscombe gives the following example (borrowed from Hare):

Do everything conducive to not having a car crash.
 Such-and-such is conducive to not having a car crash.
 Ergo, do such-and-such.

Anscombe says this premise is insane, since "there are usually a hundred different and incompatible things conducive to not having a car crash: such as, perhaps, driving into the private gateway on your left and abandoning

your car there, and driving into the private gateway on your right and abandoning the car there" (Intention, p.59). The suggestion that the major premise should be a singular statement seems plausible, because when it is expressed by a universal proposition it lacks the precision indispensable for reaching a conclusion about what is to be done, given the agent's end and the circumstances.

The minor premise of a practical syllogism is a statement about the means which have to be adopted if the end is to be realized. Kant thought that the principle of transformation of intention from ends to means is analytically true. He said "Who wills the end, wills (so far as reason has decisive influence on his action) also, the means which are indispensably necessary and in his power" (The Moral Law, p.84-5). The choice of means is determined by the agent's perception of the circumstances in which he is going to perform his action. In some cases he may come to the conclusion that there is only one means available for him. For example, if he wants a book, and the only way to get it is to borrow it from the library. But more often he has to choose between incompatible means. Suppose he wants to talk to his friend who lives in London. He can either telephone his friend, or visit him in London, or ask his friend to visit him. In cases like this, the agent may have no sufficient reason for choosing one course of action rather than another. Some philosophers suggest that the rational

agent will always choose what is the best, or what seems to him to be the best. But it is possible to argue that, as in our example, the means available to him are equally good, so he is justified in choosing any of them. We also observe that in real life people very often choose means which are not necessarily the best for the realization of their ends. From this it follows that given the agent's end and the circumstances, he may have no sufficient reason for choosing one course of action rather than another. So reasons do not always necessitate the agent's action.

It is not very clear what is the nature of the conclusion of a practical syllogism. According to Aristotle, in practical reasoning:

the two premises result in a conclusion which is an action - for example, one thinks that all men are to march and that is a man oneself: straightway one marches; or no men are to march now and that one is a man: straight way one halts. And so one acts in the two cases provided there is nothing to compel or to prevent. Again, I should make a good thing, a house is a good thing: Straightway I make a house. I need a covering, a cloak is a covering; I need a cloak. What I need I should make, I need a cloak; I should make a cloak. And the conclusion - 'I should make a cloak' - is an action. And the action goes back to the beginning or first step. If there is to be a coat, one must first have B, and if B then A, so one gets A to begin with. That the action is the conclusion is clear.

(De Motu Animalium, 701a7ff)

But Aristotle was wrong to think that the conclusion of a piece of practical reasoning is an action. For, firstly, something can prevent the performance of the action, and then it will not occur even if the premises were true. Secondly, if the conclusion is an action, then the

inference can only be drawn just before the time of the action, and it is not very clear how it can be a conclusion about a future action. Some philosophers, like Von Wright, suggest that the conclusion of a practical syllogism is the agent's declaration of an intention to perform the action. But this view has some disadvantages too. It generates a difference between the conclusion expressed in the first person by the agent involved in reasoning about his action, and the conclusion expressed in the third person by an observer. In the former, the argument issues in a declaration of an intention; in the latter, it issues in a prediction about what the agent will do. Because of these difficulties, some philosophers suggest that the conclusion should be a deontic statement which says that the agent ought to perform a certain action. Kenny, who holds this view, says "the correct account seems to be that the conclusion of a piece of practical reasoning is a description of an action to be done: a fiat concerning the reasoner's action" (Will, Freedom and Power, Blackwell, 1975, p.98). But this view generates different problems. The premises of a practical argument are descriptive, they assert what is the case; but the conclusion is supposed to be normative or prescriptive. And as Hume pointed out, one cannot draw a normative conclusion from factual premises. But perhaps this difficulty is only apparent, being generated by the word 'conclusion' which has been used in this context. Maybe the right word here is 'decision' or 'outcome'. Talk of conclusions assimilates practical reasoning too closely to theoretical reasoning.

Practical inference is entirely different from theoretical inference, i.e. a practical syllogism is not a proof syllogism. Anscombe suggests that Aristotle himself is partly responsible for this misunderstanding of the nature of the practical syllogism, because although it was perfectly clear to him that he had found a completely different form of reasoning from theoretical reasoning or proof syllogism, it pleased him to give cases of it which made it as parallel as possible to the theoretical syllogism" (Intention, p.59). The practical syllogism is not a demonstration, its purpose is not to prove that something-or-other is the case; but to show that so-and-so is to be done. The outcome of practical inference does a job which is similar in many respects to that done by explicit commands. But if so, it cannot be true or false, it is not a proposition. That is why if the state of affairs described by the outcome does not, in fact, obtain, we do not call the outcome false but we rather fault the state of affairs. To show this, Anscombe gives the following example. A man is going round a town with a shopping list in his hand. Suppose he made the list himself and it is an expression of his intention. If the list and the things the man actually buys do not agree, then, says Anscombe, "the mistake is not in the list but in the man's performance" (Intention, p.56) - (we have to assume, of course, that the man didn't change his mind in the interim). Now let us suppose that the man is followed by a detective who makes a list of the things he buys. In this case,

if there is a discrepancy between the detective's list and what the man actually buys, we will say that there is a mistake in the record, but not in the man's performance. For in this case the facts are prior, and dictate what is to be said, whereas in the former case the discrepancy doesn't impute a fault to the language but to the event, because here language functions as an order.

There is another important difference between practical and theoretical reasoning. In the latter the rules of inference ensure that we never pass from true assertions to false ones. If the premises are true and the argument is valid, the conclusion must be true. And a proposition which has the value 'true' cannot at the same time have the opposite value 'false'. But the rules of inference in practical reasoning cannot be truth-preserving because, as we already pointed out, the outcome of a practical syllogism is not a proposition but a fiat, which is neither true nor false. According to Kenny, the rules of inference in practical reasoning are satisfactoriness-preserving. He says:

In practical reasoning we never pass from a fiat which is satisfactory for a given purpose to a fiat which is unsatisfactory for that purpose. These rules are satisfactoriness-preserving, just as rules for assertoric inference are truth-preserving.

(Will, Freedom and Power, p.81)

But Kenny's suggestion cannot be right, for the following reasons. In practical reasoning the premises can have incompatible values, as in conflict situations when one has a good reason for doing one thing and at the same

time a different reason for not doing that very thing, but something inconsistent with it. For example, someone wants to write his essay, but his mother is very tired and needs his help, so he would like to help her, but it is not possible for him to do both things. Some philosophers suggest that in such cases one reason has to override the other if the agent is to act at all, so that we can distinguish between the stronger and weaker reasons. And the reason which was overridden can be regarded as not existing at all. But this is not a correct interpretation of the conflict situation. It denies that fact of human experience, that there are real conflicts of reasons. Anyway, for one reason to override another, the other has to exist, otherwise there is nothing to override. The conflicting reasons actually continue to exist even when one overrides another. And although one cannot do both actions, the reasons both apply to one's situation. But if so, then the outcome of a practical syllogism can have incompatible values, satisfaction and non-satisfaction at the same time. This means that the rules governing practical reasoning cannot be satisfactoriness-preserving. There is nothing preserved in practical reasoning, ⁱⁿ ~~that~~ the way truth is preserved in theoretical reasoning.

Practical reasoning, as opposed to theoretical, doesn't necessitate its outcome. The same premises in a practical syllogism can lead to incompatible decisions, since the agent might not have sufficient reasons for choosing one means to his end rather than

another, if they are equally good for his purposes. Suppose I want to read Plato's Republic but I haven't a copy of it. I can either buy one in a bookshop, borrow one from a library, or ask my friend who has a copy to lend it to me. Some philosophers suggest that if practical reasoning can lead to incompatible outcomes then whatever course of action the agent finally chooses, his action will be inexplicable by his reasons. But to suggest this is to deny the agent's freedom of choice between different, but equally good, means. This very freedom must lead to a loosening of the links between premises and outcome in the practical syllogism. And it doesn't matter what course of action the agent chooses, given that he had reasons for performing it, his reasons will still explain his action. But they will not explain his choice of means, for by hypothesis he had no reason for preferring any one of the alternative means available. Reasons can explain some aspects of action, but not all.

Some philosophers suggest that in a situation where the agent has only one available means to his end, reasons necessitate the action. But this is true only in some cases. In others, the agent might feel an aversion against those means, and would give up his end rather than adopt the means necessary for its realization. From this it follows that in a lot of cases the agent's reasons, given the circumstances, only incline him to perform a certain action, but don't necessitate it.

There is another kind of defeasibility involved in practical reasoning. In some cases the decision may not be acted upon at all, even if the agent didn't change his mind, and nothing prevented him from acting. In such cases, we either deny the existence of the supposed desires or intentions, i.e. the agent's reasons for acting, or we ascribe weakness of will to him. When an action is very easy to perform, but the agent didn't do it despite a declaration of his intention, we are more likely to deny that he really had that intention than to ascribe weakness of will to him. Suppose the agent said that he wanted to watch a certain programme on television, and although nothing prevented him from doing that, and he didn't seem to change his mind, he actually didn't watch the programme. In this case we would probably say that he didn't really want to see the programme. We are more likely to ascribe weakness of will when we think that the action is very difficult to perform or demands a great effort, as in the following example. The agent said that he wanted to write a paper on the nature of time, but despite this declaration he didn't write anything. In this case we are likely to believe that his failure was the result of weakness of will. However, the distinction between these two types of case is not very clear, and it is possible to have different intuitions about them.

At the beginning of the last chapter we quoted Von Wright as saying that the practical syllogism provides the social sciences with a model of explanation which can be a definite alternative to the nomological-deductive model of explanation characteristic of the natural sciences. A methodological monist may suggest that what we really presented can hardly be called a model of explanation. It provides us neither with a generalization nor with a systematic prediction, since practical reasoning is defeasible. Indeed, a practical syllogism is not a model of explanation. It is an interpretive scheme, a conceptual framework which we apply to understand an intentional action. And it is the only thing which philosophers interested in explaining human behaviour can do. As Aristotle himself said:

The theory of practice can only be schematic. Matters concerned with conduct and questions of what is good for us have no more rigidity than matters of health. The general account (of practical knowledge) being of this nature, the account of particular cases is yet more lacking in exactness; for they do not fall under any technical rule or prescription, but agents themselves in each case consider what is appropriate to the occasion to do, as happens in the art of medicine and navigation.

(Nicomachean Ethics 1104a)

Practical reasoning concerns what is ultimate and particular and constantly changing. The agent's means-deliberation system is not a closed, complete or even consistent system. It is a constant reassessment of features of a situation, and a permanent modification of the agent's own projects, interest, emotions, etc.. To attempt to subsume such a process under some

general law is to try to close and complete something which cannot be completed or closed in principle.

There is no real chance of inventing a scientific theory which can explain intentional action. Neither is there a possibility of finding some "logic" of the process of means-end deliberation. As we said before, to adopt practical principles is to be motivated to act in a certain way. But it is not true that these principles logically imply certain imperatives. The relation between ends, means and action is different from a logic of imperatives, or any other logic.

The suggestion that practical principles are governed by some such logic or can be subsumed under some general law (even if we don't know it) is a normative requirement rather than a result of investigation.

And, as Wittgenstein put it:

The requirement is now in danger of becoming empty.
— We have got on to slippery ice where there is no friction and so in a certain sense the conditions are ideal, but also, just because of that, we are unable to walk. We want to walk: so we need friction. Back to the rough ground.

(Philosophical Investigations, 107)

BIBLIOGRAPHY

- Anscombe, Intention, Oxford 1957
- _____, Causality and Determination, Cambridge 1971
- Aristotle, Metaphysics
- _____, Nichomachean Ethics
- _____, de Motu Animalium
- Braithwaite, Scientific Explanation, Cambridge 1959
- Brentano, 'The Difference between Mental and Physical Phenomena', in Realism and the Background of Phenomenology, ed. Chisholm, Glencoe 1960
- Davidson, 'Causal Relations', Journal of Philosophy 1967
- Demos, 'Self-Deception', Journal of Philosophy 1960
- Duhem, The Aim and Structure of Physical Theory, Paris 1914
- Feigl, 'The Mental and the Physical', in Minnesota Studies in the Philosophy of Science, Vol.II, edited by Feigl, Scriven and Maxwell, Minneapolis 1958
- Hacker, Insight and Illusion, Oxford 1972
- Hart and Honore, Causation in the Law, Oxford 1959
- Hempel, 'Explanation in Science and History', in Frontiers of Science and Philosophy, ed. Colodny, Pittsburgh 1962
- Kant, Foundations of the Metaphysic of Morals
- Kenny, Will, Freedom and Power, Oxford 1975
- Kim, 'On the Psycho-physical Identity Theory', American Philosophical Quarterly 1966
- _____, 'Causes and Events: Mackie on Causation', Journal of Philosophy 1971
- _____, 'Causes and Counterfactuals', Journal of Philosophy 1973
- Kripke, 'Naming and Necessity', in Semantics of Natural Language, edited by Davidson and Harman, Dordrecht 1972
- _____, 'Identity and Necessity', in Identity and Individuation, edited by Munitz, New York 1971
- McGinn, 'Anomalous Monism and Kripke's Cartesian Intuitions', Analysis 1976-7
- Mackie, 'Causes and Conditions', American Philosophical Quarterly 1965
- Melden, Free Action, London 1961
- Pears, 'Desires as Causes of Action', in Questions in the Philosophy of Mind, London 1975
- _____, 'Rational Explanation of Action and Psychological Determinism', in Essays on Freedom of Action, ed. Honderich, London 1973

Spinoza, Ethics

Toulmin, Foresight and Understanding, London 1961

Wilkes, Physicalism, London 1978

Wittgenstein, Philosophical Investigations, Oxford 1953

von Wright, Explanation and Understanding, London 1971