THE ROLE OF CAUSES IN ACTION EXPLANATIONS

TWO COMPETING APPROACHES

Haralabos Hatziioannou

Mphil in Philosophy

LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

April 1997

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

The thesis deals with the question of whether causation can play a (relevant) part in the explanation of action; it approaches it through the critical assessment of two paradigmatic theories of action, one of each side of the debate. In the first part D.Davidson's Causal Theory of Action is presented (as a development of the Causal Nomological Theory), and criticised on the grounds that it cannot provide adequate singular causal explanations of actions. The argument questions Davidson's model by challenging the way in which Davidson could justify the causal relevance of reasons. It concludes that the causal condition adds no additional explanatory force on non-causal rationalisations. In the second part, the nature of action explanations is examined through von Wright's non-causal approach. After presentation of his theory, his version of the Logical Connection Argument is considered in the light of various criticisms that have been directed against it. Although his argument is found to be inconclusive with respect to the impossibility of a Causal Theory of Action, it is argued that, the implications that follow from von Wright's discussion of it, render the causal claim irrelevant to the explanation of action. Finally von Wright's possible response to two types of criticism, is considered: The justification Vs explanation argument and the problem of congruence. The conclusion is that his theory deals in a satisfactory way with the first one, but fails to meet the challenge of the second one.

To Christina, my first teacher

CONTENTS

Preface

In this thesis, the question of the role of causes in the explanation of human action, is dealt with, through the examination of the views of two philosophers which take opposite sides on the debate. D. Davidson's theory of action is initially presented as a natural development of the Causal Theory in its original Nomological form, but, as I will try to show in the subsequent discussion, Davidson cannot sustain all of his central claims in the Philosophy of Psychology, without making any concessions. My inclination is to question the relevance of causation in the explanation of action, rather than Davidson's views on the impossibility of formulating strict laws in psychological terms, and so, in the second part, I present and assess von Wright's intentionalist approach, which is probably the most comprehensive and influential non-causal theory.

The juxtaposition and assessment of the two theories, is not meant to provide a final answer to the central question that I consider. I am sure that a lot more can be said in defence of, or against either theory. My aim is rather to show how consideration of the arguments presented by the two philosophers can provide some ground to question the relevance of causation in the explanation of action. These considerations are not decisive since, as I will argue in the second part, no version of the Logical Connection Argument can be conclusive and, furthermore, von Wright's theory faces some serious problems of its own.

As will become apparent from the discussion, I do not think that everything which is of importance to my central question can be said within the context of these two theories. However, I have decided to concentrate mainly on them, not only because of their completeness and their decisive influence in the development of the main themes in the Philosophy of Action, but also for the purpose of keeping a fixed frame within which my main concern can be addressed. There is certainly a multitude of interesting views that are left out and which may be helpful in dealing with the problem I am addressing; hence, the role of causation in the explanation of actions must remain an open issue for me.

ACKNOWLEDGEMENTS

There is a great number of people and institutions without whom the completion of my thesis would have been impossible. I would like to express my gratitude to the Onassis Foundation for providing me with funding during the second year of my studies, as well as the Philosophy department at L.S.E., which awarded me the Lakatos Scholarship during the first two years. All the members of staff at the L.S.E. department have been especially supportive and helpful in the eventful course of my studies there. Especially, I want to thank my supervisor, Professor D.-H. Ruben who started from scratch with me, and has helped me reach a level in Philosophy which would have been much lower, without his guidance and instruction. Other individuals from the L.S.E., whom I would like to thank for their help, are Demetris Portides, George Zouros and especially George Keranis who has provided advice and criticism at almost every step of the project. I am also grateful to Andigone Kyriakides and Olatz Ugarte for their valuable contribution to the completion of this thesis. Finally, my greatest indebtedness is owed to my mother, my brother, my sister and my late father, for, without them I would not have even started this project.

H. Hatziioannou, Athens March 1997
Scholar of the Onassis Foundation

PART I: THE CAUSAL APPROACH

CHAPTER 1: THE CAUSAL-NOMOLOGICAL THEORY OF ACTION

Statement of the Theory

The Causal Theory of Action is a theory that construes explanation of behaviour as a special kind of causal explanation, not differing in principle from explanation in the Natural Sciences. This claim is meant to counter a quite widespread and persistent view, which holds that the explanation of human action in terms of purpose is of a fundamentally different kind, conforming to the model of teleological explanation and not admitting the use of the concept of efficient causation. To explain a particular action X, is according to this view, to state the agent's goal, Y, say, and his belief that by doing X he would bring about Y.

C.J. Ducasse, one of the first modern proponents of the causal theory, argued[1] that such explanations could, and should be construed in causal terms. The desire of the agent that Y shall occur and his belief that by doing X he will bring about Y, can be thought of as joint causes of the agent's behaviour in doing X, in the same sense of the word 'cause' as that used in any scientific causal explanation. According to Ducasse, this kind of explanation 'essentially consists in the offering of a hypothesis of fact, standing to the fact to be explained as case of antecedent to case of subsequent of some already known law of connection '[2]. Hence, the central idea of the Causal Theory of Action in this form, is that the explanation of action consists in stating a causal law ('laws of bare conjunction statistically obtained, will not do'[3]), and a set of particular facts (the agent's desires and beliefs and the relevant conditions obtaining), from which the behaviour to be explained follows as a consequence.

---

[1] In Ducasse, [1925].

[2] Ibid. p. 151.

[3] Ibid. p.151.

This idea found its most natural and elaborate expression in Hempel's model of Deductive-Nomological explanation. Hempel developed his model for explanation by nomic subsumption in a systematic and precise way, claiming for it the status of an ideal standard for *all* complete, scientific explanations. In particular, he argued that the logic of action-explanations does not differ from that of any other genuine, (as opposed to 'pseudo') scientific explanation and he offered a model, in accordance with the D-N, which took into account the peculiarities characteristic of action-explanations. Although he did not place emphasis on the idea of *causation* as the explanatory relation between reasons and actions, since he considered it of secondary importance to that of lawful connection, he certainly allowed and accounted for it in terms of causal laws. Therefore, it is right to say, that the Causal-Nomological Theory is best developed as a theory that construes action-explanations in accordance with Hempel's covering-law model.

According to Hempel, a complete explanation of an event of kind E, consists of:

> '(1) a set of statements asserting the occurrence of certain events C1.....Cn at certain times and places,
>
> (2) a set of universal hypotheses, such that
>
> (a) the statements of both groups are reasonably well confirmed by empirical evidence
>
> (b) from the two groups of statements the sentence asserting the occurrence of event E can be logically deduced.'[4]

It should be noted that the explanandum is a statement asserting the occurrence of a concrete event of a *specified kind* (E). It is not claimed that such an explanation is complete in the sense that it accounts for all properties that the concrete event, **e** say, possesses; the occurrence of **e** cannot be explained without reference to *any* of its characteristics. It follows that the deduction of the explanandum from the explanans is possible, only if the laws featured in the latter make explicit mention of the property by which the concrete event **e,** which features in the explanandum, is picked out, i.e. its property of being of kind E.

Action-explanations in terms of motivating reasons, besides the usual features of all D-N explanations, have the special characteristic that they are connected with the normative idea of rationality. For Hempel, to give an action-explanation in

terms of reasons, involves commitment to two claims: First that the action was *motivated* by these reasons (which is an empirical hypothesis) and, second, that the action explained is *appraised* as being the appropriate, or reasonable thing to do, given the agent's reasons and the particular circumstances obtaining (which is a critical, or normative appraisal). These conditions are satisfied by a D-N action explanation, in view of the inclusion in the premises of the D-N argument of a general law, which states that all rational agents act in a specified way under a certain set of circumstances, and of a statement of particular fact to the effect that the agent under consideration was (at the time of the action) a rational agent. Of course a further premise, stating that the conditions specified in the antecedent of the general law were holding at the relevant time, is needed for the deduction of the explanandum to be possible. Accordingly, Hempel construes his D-N model for action explanations as follows:

'A was in situation of type C
A was a rational agent
In a situation of type C, any rational agent will do x
Therefore, A did x'[5].

The first premise, which specifies the agent's condition before the action, will include ascriptions of beliefs, desires and other attitudes that may be influential to the agent, as well as all the relevant to the performance of the action facts. Together with the second premise, they represent the singular statements of fact that are necessary for every covering-law explanation to be deductive. The third premise is not, according to Hempel, a mere 'principle of action' which expresses only the *appropriateness* of actions of type x in circumstances of type C, but an *empirical law*[6] stating, in broadly dispositional terms, how rational agents with the kind of beliefs and desires specified in the first premise behave. As an empirical

---

[4] Hempel [1942], p.232.

[5] In Hempel, [1965], p.471.

[6] Hempel produces an analysis of dispositional properties in terms of their various symptoms (or manifestations) in different circumstances, in defence of his thesis that such laws are empirical, rather than merely analytic (see e.g. Hempel [1965], pp.457-63).

law, it performs the main explanatory function among the explanantia: It tells us why, given that the first two premises were true at the time of action, the agent did, in fact, do x. If we knew the initial conditions and the general law before the action, we could have predicted with certainty that the agent would do x. If no such law was given among the explanantia, then the deduction from the premises to the conclusion would not go through, and thus we would have no reason, given the initial conditions, to expect that the agent would perform the action.

Hempel, with his D-N model, purports to provide a reconstruction of action-explanations that makes their logical structure, and thus their explanatory features, transparent. In actual scientific practice (in History and the Social Sciences) such complete explanations are seldom, if ever, given. The explanations that are actually given are just explanation-sketches in the sense that they are either *elliptical*, or *partial* (or both). Elliptical explanations do not make explicit mention of the general laws or of some of the particular facts involved in the explanation but, if they are to be considered genuinely explanatory, they must be thought of as tacitly assuming the missing premise(s) of the argument. In this case, it must be possible from the context of the explanation, to extract and formulate precisely the law or fact missing, thus turning the explanation into a complete one.

In the case of partial explanations however, it is not possible to give the exact formulation of the law that is imprecisely stated in the premises, or missing. This may be so because some of the relevant characteristics of the particular facts of the explanans or the explanandum are either missing, or are not stated with sufficient specificity so that, they can be brought 'in contact' with all the empirical evidence available and the general law be inferred. Partial explanations are not, strictly speaking, D-N explanations, since the explanandum cannot be logically deduced from the premises. They are just explanation sketches, which need filling out of their missing or imprecise terms in order to be turned into full explanations. The direction that our investigation into the missing elements will take, is usually suggested by the empirical facts and the concepts that already feature in the explanation sketch. This process is empirical and depends crucially on whether the explanation sketch on which it is based, is genuinely explanatory or only apparently so. For, an explanation sketch which is based on scientifically acceptable empirical evidence and theories, will indicate the ways in which its

precision can be increased and the relevant factors specified. On the other hand, 'pseudo'-explanation sketches are, so to say, 'sterile', since they give no clues as to what amendments and additions they need, in order to be turned into full explanations with empirically testable explanantia.

The Methodological Criticisms

Hempel construed as a necessary condition for the adequacy of an action-explanation sketch that it makes at least implicit reference to a law-sketch that can, with further empirical enquiry, be turned into a full-blown law. However, even this 'watered-down' requirement for adequacy is rarely met by actual scientific explanations of human action. This is a problem that, by itself, is not sufficient to render the model inadequate as an exposition of the *logical structure* of such explanations. Hempel's theory of explanation provides us with a schema that is supposed to give us insights into what the notion of explanation amounts to, and set an ideal standard for which all explanations must strive; it does not simply *describe* the actual process of explanation as it is performed by scientists or laymen.

However, W. Dray[7], argued that the discrepancy that exists between 'ideal' covering-law explanations or even explanation sketches of the kind specified above, and explanations as they are actually given by historians (or, for that matter, by social scientists and psychologists), is due to the fact that the model fails to capture the nature of the concept of explanation in these sciences. He offered negative arguments based on methodological considerations, as well as a positive theory of the model of 'rational explanation', which he thought to be peculiar to human action.

Dray's arguments do not bear directly on the logic of action-explanation as this is laid out by Hempel's model, and so do not amount to *a refutation* of the model.

[7] See Dray [1957] and [1963].

The standpoint from which Dray argues is the methodological one, claiming that Hempel's theory is methodologically fruitless for explanation in the Social Sciences. Explanations that are considered as adequate by practitioners in these fields, *need not*, and in fact *do not* refer to general laws. To consider them as explanation-sketches, which point the way to the development of full-fledged covering-laws, is fruitless from the methodological point of view, since any laws that might be proposed, will most probably be either vacuous or of a very low level of generality; in either case they will confer no additional force to the already existing explanation in terms of the agent's reasons. Serious, explanatory causal laws are just not suggested by explanations that are actually given by social scientists and historians. To insist that *some* covering law must be tacitly assumed by *any* adequate explanation in these fields, Dray argued, is to miss the *normative* element, which is the most important feature of such explanations. The objective of the historian, for example, in explaining the actions of historical figures, is to show them as being rational in the light of the agent's conjectured aims and beliefs.. To state a general hypothesis of the form 'all rational agents in such and such circumstances, act in such and such a way' does not amount to a demonstration of the action's *rationality* in the *particular conditions* obtaining. Such a regularity does not help us in our appraisal of the action in its particularity.

Although Dray's arguments make a strong case against the model, his own positive account of the explanation of action is not fully developed, and has raised cogent criticisms from causal theorists. Hempel[8], in his own reply to Dray's criticisms, is mainly concerned with showing that, unless his model is accepted, there can be no standard of complete explanation of action; an account such as Dray's, only succeeds in showing why an action was, from the agent's point of view, a reasonable thing to do, but not why the agent did, in fact, perform the action. On Dray's account, we can show that the agent had reasons for acting as he did, but we cannot show why he *did* act that way. Only with the additional covering-law premise can an action-explanation be complete, in the sense of closing the entailment gap, and allowing us to assert that, had we known the explanantia before the action, we would have predicted with certainty that it would

[8] e.g. in Hempel [1963] and [1965].

follow. Hempel's criticism against Dray has been a standard criticism of causalists, in defence of their thesis that only a causal theory can provide complete explanations of human actions. The criticism is cogent as far as the inadequacy of Dray's account is concerned; but it does not show that the causal theory of action-explanations *must* be correct. In particular, Dray's arguments against the Causal-Nomological Theory are sufficient to show its implausibility as far as its construal of action-explanations is concerned.

CHAPTER 2: DAVIDSON'S CAUSAL THEORY OF ACTION

D. Davidson, with his Causal Theory of Action, attempts to deny the causalist's commitment to covering-law explanations, while retaining the claim that human action affords a causal analysis. Davidson recognised the force of the arguments against the Causal-Nomological Theory, and proposed an analysis of action in terms of event causation, without direct recourse to nomic generalisations to support the causal connection between reasons and actions. His solution is based on a simple and elegant ontology of events and an ingenious construal of the logical form of singular causal and action sentences. It is also supplemented by a theory of 'non-reductive' materialism, the theory of Anomalous Monism, whose main tenets were already suggested in his influential first paper on the theory of action[9], but was fully developed and defended in his subsequent works[10]. Davidson's contribution to the Philosophy of Action, has triggered off the revival of the Causal Theory and has itself come to occupy a distinct and influential position in the spectrum of theories of action. However, it has also drawn some powerful arguments against it, the most important of which, I will try to modify and present in the next chapter. But first, I will give an outline of the essentials of Davidson's theories. Although the question of the analysis of action is inextricably linked with the question of its explanation, I will attempt here to separate Davidson's answers to the two questions, since my arguments in the next chapter will be directed against his model for the explanation of action, bearing only indirectly on his analysis.

The Analysis of Action

---

[9] 'Actions, Reasons and Causes', in Davidson [1980], pp.3-19. (Hereafter ARC)

[10] Mainly in 'Mental Events', in Davidson [1980], pp.207-27 (Hereafter ME)

With his theory, Davidson purports to provide an analysis of action in terms of events and the relation of event causation. This analysis is not meant to be a *conceptual* one, i.e. it does not claim to analyse what we *mean* when we use the concepts of agency and action[11]. Our notions of agency and action may be thought as basic as that of causation, but they might still be just ways of talking, involving no commitment to basic entities such as actions and agents, or relations such as that of agent-causation. Thus, Davidson's theory is conceived as an ontological analysis of action, admitting as ontologically basic only events which are causally related. To be sure, in the course of arguing for his theory, Davidson appeals to our ordinary notions, but he does this only to show that, if his theory is correct, then they need not commit us to the existence of any other entities besides events, or any other relation besides the causal one.

Events, in Davidson's theory, are thought to be concrete particulars, referred to by singular statements such as 'the death of Caesar', 'the first explosion of atomic bomb in war history', 'John's twentieth birthday'. Events can take different true descriptions as, for example, 'the first explosion of atomic bomb in war history', can also be described as 'the explosion at Hiroshima on 6th August 1945'; both descriptions refer to the same concrete event, whose existence is independent of any linguistic expression of it.

The causal relation is understood to be a contingent dependency relation, whose relata are events. So, a singular causal statement of the form 'c causes e', reports an extensional relation between the concrete particulars c and e; if 'c causes e' is a true causal statement, we can substitute, *salva veritate*, for c and e, any of their logical equivalents or any singular terms referring to them. For example, the truth of the singular causal statement, 'the short-circuit caused the fire', is not affected by replacement of 'the fire' by some other true description of the event referred to. So, if the first statement is true, we can also truly assert that, 'the short-circuit caused the most amazing spectacle I have ever seen', provided that 'the most amazing...' is a true description of the event picked out by the term 'the fire'. Davidson does not attempt to provide an analysis of causation, but only an analysis

---

[11] This position follows from Davidson's arguments in 'Agency', in Davidson [1980] (hereafter EAE), pp.43-61

of the *logical form* of singular causal statements, claiming that the two tasks are independent. His arguments[12], yield the conclusion that a singular causal sentence does not entail any *particular* law that connects the cause and the effect under the descriptions used in the sentence. What they entail, is that *some* law, that uses some appropriate true descriptions of the events involved, exists. We can pick out the events related as cause and effect, without using, or suggesting, the appropriate descriptions under which they are lawfully related: Knowing that a singular causal statement is true, does not entail knowing, or even having some idea of, the law that subsumes the events related. For example, one may know that 'the short-circuit caused the fire', even if he has no idea of the law that 'grounds' this causal relation, and even if the law describes the cause and the effect using entirely different terms than 'short-circuit', and 'fire'.

Using the extensionality of the causal relation, Davidson was able to retain the causal thesis, while rejecting Hempel's covering-law model for the explanation of action. He claimed that the agent's reasons may still be causes of his actions, even if there are no causal laws connecting the corresponding events when described as reasons and actions. So, according to his analysis, intentional action consists in behaviour which is caused by the onset of a complex mental state of the agent. This mental state he called a *primary reason* for the action, and took as its necessary constituents a pro-attitude of the agent towards actions of a specified kind, and a belief of the agent that the action, under the description believed by him to be true, is of this specified kind.

The first thing to note here is the ambiguity of the locution 'consists in'. Davidson advanced his analysis of intentional action as an analysis in terms of necessary, but not sufficient conditions, thinking that a *full* analysis in terms of necessary *and* sufficient conditions would subsequently become possible[13]. But, in

---

[12] Developed mainly in his 'The Logical Form of Action Sentences' and 'Causal Relations', in 'EAE', pp.105-22 and 149-62 respectively.

[13] As, for example the concluding sentences from 'ARC', p.19, suggest: ' Some causes have no agents. Among these agentleses causes are the states and changes of state in persons which, because, they are reasons as well as causes, *constitute* certain events free and intentional actions'.

his later papers, he came to believe that a full analysis was not, after all, possible[14]. What forced him to this pessimistic for his theory view, was consideration of the problem of so-called deviant, or wayward causal chains. Cases involving deviant causal chains were introduced as counterexamples to the Causal Theory of Action, by R.Chisholm[15], and were recognised by Davidson as presenting a real problem for it. They show that the causal analysis of action proposed is too broad to capture exactly the class of actions, since it counts as intentional action behaviour that we would intuitively recognise as non-intentional. The example given by Davidson himself, is that of a climber who 'might want to rid himself of the weight and danger of holding another man on a rope, and he might know that by loosening his hold on the rope he could rid himself of the weight and danger. This belief and want might so unnerve him as to cause him to loosen his hold, and yet it might be the case that he never *chose* to loosen his hold, nor did he do it intentionally'.[16]

In this example, we have a case where all of Davidson's conditions are fulfilled, since the climber's want and belief constitute a primary reason, which is also a cause of his behaviour. Nevertheless, our intuition strongly suggests that the climber's loosening of his hold is *not* an intentional action. The problem is that the climber's primary reason did not cause the behaviour *in the right sort of way*, in a way that is, that would allow us to say that he acted intentionally. The reason caused the behaviour via a deviant path, in the same sort of way that fearful thoughts say, may cause someone to perspire or tremble. Perspiration and trembling can have volitional and cognitive states as causes, but they certainly are not actions, but reflexes. Similarly, the climber's loosening of his hold on the rope cannot be classified as intentional action, as Davidson's theory would seem to require if we accepted his conditions as *sufficient*.

The problem is that Davidson cannot add a clause to the effect that the agent's reasons must cause the behaviour *in the right sort of way*, since the only thing that one can mean by that, is that the agent *acted intentionally*. Hence, such an addition

---

[14]See for example his 'Freedom to Act', in EAE', p.80: 'We must count our search for a causal analysis of «A is free to do x» a failure.'

[15] See his [1966], paper.

[16] 'Freedom to Act', in 'EAE', p.79.

would not help Davidson, since it would appeal to the very notion that he set out to analyse, that of intentionality.

The problem of deviant causal chains has been a standard theme in the literature on the Philosophy of Action, and many attempts have been made to treat it. It will not be my main concern here, since I think that the argument that I will present against Davidson' theory will show it to be a symptom of a deeper problem that the theory faces. However, Davidson's admittance that a solution is not forthcoming, raises a question as to how we should understand his position. If a *full* analysis of intentional action is not possible, then he cannot claim that the 'embarrassing entity' that it represents has been removed completely from the 'world's furniture', for it is still needed to distinguish between the deviant and the non-deviant cases of causation of behaviour by reasons. Hence, without necessary *and* sufficient conditions for the analysis of intentional action, 'actionist' residues *must* remain in the 'world's furniture', and thus the Causal Theory of Action is indefensible as a theory of ontological reduction of intentional action.[17] In view of these considerations, Davidson's position seems puzzling to me. This is why I will consider his theory as an attempt to provide a *full* analysis of action, and argue against it without using the counterexamples of deviant causal chains. As said before, I believe that there is an underlying problem that will become apparent after examination of his model for the explanation of action. Hence, for the moment, I will set aside the problem presented by deviant causal chains, and treat Davidson's analysis as a *full* ontological analysis of intentional action.

I have reconstructed Davidson's analysis as an analysis of intentional action, rather than simply of action, since, for him, the latter is secondary to the former. An agent may act non-intentionally only if under some other true description, he is acting intentionally. If, for example, I mistakenly turn on the ventilator, thinking that the switch on the wall was the light switch, I have intentionally done *something*, namely, flipped the switch. For Davidson an action is an event (with the right sort of causes), that can take on different descriptions: Under some of these descriptions it is intentional, under some others, it is not. But for an event to be an action at all, there must be a description that makes it an *intentional* action.

---

[17] For a similar view on Davidson's position ,see Bishop [1989], pp.101-5.

Davidson takes the 'extreme minimising' view on the issue of the individuation of action. This is brought out by the term 'accordion effect' which characterises the way we can shorten the description of an action down to the agent's bodily movements, or, we can stretch it to its furthest consequences. However, despite all this contracting and expanding, the action that the agent performs remains the same; it is only the descriptions in terms of its causes or effects that change. If for example A wakes up her husband who is sleeping in the room by turning on the light, the only event of which A is the agent, is the movement of her hand. All other descriptions, like e.g. 'flipping the switch', 'turning on the light', 'illuminating the room', 'waking up her sleeping husband', are descriptions of the same primitive (basic) action in terms of its effects. Under some of these descriptions, A's action is intentional (e.g. 'turning on the light'), under some others it is unintentional (e.g. waking up her husband). Hence, Davidson's view amounts to saying that all actions are, ultimately, bodily movements. The richness with which we can describe actions, is due to the multitude and variety of the effects of agents' movements; behind it, there is only the bodily movement itself.

The question of the individuation of action is very important in itself and has given rise to great controversies between philosophers of action. In my treatment of Davidson's theory, I will not be concerned with it, taking his 'extreme minimising' view for granted. So, whenever I speak about actions, I will mean the descriptions of some intentional bodily behaviour in terms of its effects (or circumstances); the *descriptum* will always be a single event: the agent's intentional primitive action.

Davidson's Model for Singular Action Explanations

Let us examine now the model for singular explanation of action which is proposed by Davidson.[18] According to it, there are three necessary conditions for an adequate explanation of action:

[18] This is a reconstruction of Davidson's views on the explanation of action which follows, to an extent, Antony's ([1989]). Davidson, as far as I know, has never presented his views in this way. The justification for my reconstruction will become, I believe, apparent, with the qualifying remarks that follow.

1. Any attribution of mental attitudes to the agent contained in the explanation, must be true.(The Truth Condition).

2. The explanation must attribute to the agent, in accordance with the Truth Condition, a primary reason, consisting of a certain pro-attitude and belief, that display in their light the action as being reasonable. (The Rationality Condition).

3. The advent of the primary reason featuring in the explanans, must be the cause of the action. (The Causal Condition).

Let us look at the role that each condition plays in the explanation of action:

1. The Truth Condition is necessary if we are to take a realist stance towards the propositional attitudes that are supposed to populate the mind of the agent, providing the content of his reasons for action. By propositional or intentional attitudes, we mean mental states that are expressed by verbs like believing, desiring, expecting, intending, knowing, and so on, and which exhibit the characteristic of being directed at, or being about, existing, or non-existing objects or states of affairs. Their ascription to agents creates non-extensional contexts, i.e. sentences whose truth-value may change after substitution of one of their terms by one of its logical equivalents, or by a term which is co-referrent with it. For example, the sentence, 'A knows that the capital of Chile is in South America', may be true, with the sentence, 'A knows that Santiago is in South America', being false.

It is not clear to me, what is the exact treatment of the intentionality of mental states by Davidson, but, at least, this much is certain: That he is a realist about events that can (truly) be described as the advent of a belief, a desire, a want, or, more generally, a pro attitude. That is, when we describe an event **e**, as the coming to believe that **b** by agent **A**, we are giving a true description of a real entity, namely of **e**. In this sense, Davidson is a realist about propositional attitudes, even though he is not committed to the existence of intentional states or properties (or, for that matter, to the existence of any state or property). The question of how he accounts for the intentionality of these descriptions, of how, that is, they are connected to the real entities in the world, is a separate issue, which I will not take up here.

So, the Truth Condition is a necessary condition for taking a realist stance towards the propositional attitudes, which, in turn, is necessary for rationalisation

being a type of causal explanation. For, if Davidson was not a realist about the attribution of desires and beliefs, then he could barely explain an action by reference to them. There has to be some sense in which it is true to say of an agent **A** that she came to believe that **b**, in order to be in a position to claim that **A**'s belief that **b**, was part of the cause of the action. If the existence of the belief is disputed by some theorist, then he cannot place that belief in the explanans of the action, unless he is a non-realist about explanation, which Davidson clearly is not. Davidson's way of being a realist about propositional attitudes, is by accepting that the intentional descriptions of the concrete events, are true descriptions of real entities: to describe an event **e** as the advent of **A**'s belief that **b**, is therefore, to give a true description of a concrete particular[19].

2. The Rationality Condition is accepted, with certain modifications, by causalists and non-causalists alike, as a necessary condition for the explanation of action. It is supposed to ensure 'that the agent is shown in his role of Rational Animal'[20]. In order to do this, we have to be able to construct a practical syllogism, with premises the desire and the belief of the agent, and conclusion the assertion that the action had some 'desirability characteristic' for him. Davidson does not accept that practical syllogisms are necessarily involved in practical reasoning. Rather, it is part of the analysis of our concept of a reason for which an agent acted, that it is possible that such a syllogism be constructed. It is not necessary that the logical (or psychological) steps of the argument are actually followed in the process of forming an intention to act. It is only necessary that the reason for the action possessed by the agent, allows for such a practical syllogism to be constructed.

Thus, a primary reason justifies an action only in a weak sense, showing it to have a property which is desirable for the agent. It does not justify it in the strong

---

[19] Whether reasons can be taken as causes of action has been the point of an objection raised by R.Stoecker, on the grounds that reasons are not events. (See his [1993] paper) This is an interesting issue, which I will not take up here, especially in view of Davidson's insistence that the *advent* of a reason can be considered as the cause of an action. (See his 'Reply to Stoecker' in the same volume.). In what follows, whenever I refer to a *reason* as the *cause* of an action, I will mean, as Davidson does, the *advent* of that reason.

[20] ARC, in EAE, p.8.

sense of the word, i.e. showing it to be desirable *in itself.* Still, it satisfies our demand to read rationality in the action, to see it, that is, as a piece of behaviour which is coherent with a certain trait, preference or desire of the agent. If we were not in a position to produce even a single property of the action that seemed appealing to the agent, then the action would be unexplainable, indeed it could not be characterised as action at all.

Note that a primary reason, being a pair of intentional states, can explain an action, only insofar as the action is described in a way that is known to the agent. For example, we cannot explain Oedipus' striking of his father by his desire to punish the rude old man that was found in his way. Although the two descriptions ('striking the rude old man' and 'striking his father'), apply to the same action (event), the latter is explained only when described as 'a striking of a rude old man'. Only the descriptions which are *thought* by the agent *to be true*, not *the* true descriptions, matter in explanation. The intensionality of mental attitudes, infects the whole action explanation.

3. The Causal Condition must be included as a necessary condition in the analysis of action explanation by any causal theory of action. Davidson proposed it as a necessary condition, supporting it with an argument against any theory that does not include such a condition in it. How can we distinguish, Davidson asked, between an agent's acting *because of* her reasons, and her acting and *merely having* these reasons? An agent may have adequate reasons for an action, but nevertheless perform it for a different set of (adequate) reasons that she also possessed at the time of the action. Consider for example the following case, where there is a clear distinction between: (i) She exercised *and* she wanted to reduce weight, and thought exercise would do it, and (ii), She exercised *because* she wanted to reduce weight, and thought exercise would do it, which cannot be accounted for by explanation models that contain only the Truth and the Rationality Conditions.

Thus, merely citing the reasons that the agent possessed at the time of action, is not enough for explanation. At most it may be enough for justification. Therefore, a further condition is needed to supplement the first two, so that justification is turned into explanation. Explanation by non-causal rationalisation is, Davidson argued, a kind of explanation which is not well understood, and which the non-

causal theorists had not defended adequately, by explaining how it works. Causal explanation, on the other hand, is a perfectly legitimate kind of explanation, a kind which is understood as well as any. So, Davidson concluded, the only viable proposal for accounting for the explanatory force of 'because' in (ii), is the causal one. Causation guarantees that there exists a real connection between reasons and action: The onset of a reason is the cause (or a part of the cause) of the action, and this is enough to ensure a link between the two, a link that makes reasons explanatorily relevant in a well understood way.

Hence, Davidson views rationalisation as a kind of causal explanation of a special class of events that are called actions. It is certain that he regards rationalisations as adequate explanations, but whether he thinks they are complete is not clear, since he is not trying to characterise explanation generally and thus make plain what he would take to be a complete explanation. He writes:

> ' We may join in lauding as an ideal explanation a description of antecedents and a specification of laws such that the explanandum can be deduced; but how much less still counts as explanation? It seems to me that we have in action a particularly good specimen for study; since we agree that one way of explaining actions is by giving the agent's reasons, we can concentrate on the relatively clear question what reason explanations are like, and set aside the more diffuse problem of characterising explanation generally.'[21]

Anomalous Monism

Davidson's Causal Theory of Action is supplemented by a Theory of Mind which he calls Anomalous Monism. In it, he combines three apparently inconsistent principles, by endorsing a 'token-token' Identity Thesis. The three principles are stated as follows:

1. The Principle of Causal Interaction: 'At least some mental events interact causally with physical events'.

[21] In 'Hempel on Explaining Action', EAE, p.263.

2. The Principle of the Nomological Character of Causality: 'Events related as cause and effect fall under strict deterministic laws'.

3. The Principle of the Anomalism of the Mental: 'There are no strict, deterministic laws, on the basis of which mental events can be predicted and explained'.[22]

The apparent contradiction is that the first two principles appear to entail the denial of the third: If there are mental events that are causally related to physical events, then, by the second principle, they must fall under strict laws; but, by the third principle, there are no laws that can be used to predict and explain mental events.[23] Davidson escapes the seeming contradiction by affirming the thesis of Anomalous Monism, while asserting the extensionality of singular causal statements and the linguistic (and hence intensional) character of laws. Anomalous Monism is the claim that every mental event is *token-identical* to some physical event, but mental *types* are not identical, or reducible to physical *types*. Davidson's understanding of singular causal sentences, has already been presented: Causal relations hold between events; events are particulars that can take on an indefinite number of descriptions, and only under *some* of these description, they can instantiate laws. Hence, laws are linguistic, relating events only insofar as the latter are described one way or another.

Thus, the apparent contradiction is resolved as follows: An event which has a mental description, and which is causally connected to some physical event (as, according to the first principle, at least some mental events are), is subsumed under the strict law that (according to the second principle) *must* exist covering the connection, only when physically described. Hence, the third principle is not violated, since no mental descriptions of events feature in strict laws.

Most aspects of Anomalous Monism are already either explicit or implicit in his Causal Theory of Action. The Principle of Causal Interaction has a clear manifestation in actions: Actions are physical events (bodily movements), which

---

[22] In his 'Mental Events', in 'EAE', pp.207-27

[23] As stated, Davidson's third principle is not strictly correct: Mental events *can* be predicted and explained using strict laws, since every mental event is identical to some physical event, and physical events can be explained and described by strict laws. Hence, the principle must be understood as saying that there are no strict laws that contain mental *descriptions* of events.

are causally related to mental events (the onsets of reasons). Furthermore, his claim that the Causal-Nomological Theory of Action was not the only causal alternative involved accepting the Principle of the Nomological Character of Causality, but understanding the logical form of singular causal statements along the lines sketched out above. Finally, the token-token identity thesis, also has a manifestation in his Causal Theory of Action: Reasons and actions are events that can take both a mental (intentional) and a physical description. Of course, in 'Actions Reasons and Causes', Davidson did not go all the way to outright denial of the possibility of strict psychological and psychophysical laws, but his thesis that reasons *need not* be lawfully related to actions, suggests at least some scepticism as to whether they *could* be so related.

Thus, Anomalous Monism apparently fits well with Davidson's Causal Theory of Action: The causal connection between reasons and action could be thought of as the paradigmatic case where all the principles on which Anomalous Monism is based, are instantiated without contradiction. However, Davidson's emphatic denial[24] of the possibility of formulating strict laws featuring mental terms has, according to some critics, serious consequences for his Theory of Mind. In the next chapter, I will present a version of this type of criticism, and assess its bearing on Davidson's theory of action.

[24] His arguments are expounded in his 'Mental Events' as well as 'Psychology as Philosophy', in 'EAE'.

CHAPTER 3: A CRITICISM OF THE THEORY

Introduction

D.Davidson's Anomalous Monism has been widely criticised[25] for rendering the mental causally inert. The objection is based on the fact that, in Davidson's theory, events that enter into causal relations instantiate the laws that subsume these relations, only when physically described. Davidson's token identity thesis, the claim, that is, that the mental and the physical descriptions both apply to the same entities while mental types are irreducible to physical types, does not save, according to the critics, the theory from the charge of Epiphenomenalism. They argue that, since the causal relations involving mental events are backed only by physical laws which make use of physical descriptions of the events related, then the latter are causes and effects only *in virtue of* their physical properties and not *in virtue of* their mental properties. Another way of putting it, is to say that events enter into causal relations *qua* physical and not *qua* mental events. Thus, the critics conclude, Davidson's Anomalous Monism, renders the mental causally inert and hence the theory collapses to a type of Epiphenomenalism.

The consequences for Davidson's Causal Theory of Action are no less serious, since, if the criticism is correct, reasons are not causes of actions in virtue of their intentional characteristics (i.e. their content), but in virtue of their physical ones. The fact that a person's belief is a belief *that p,* and that her desire is a desire *to q*, will be as causally relevant to her action as is the brick's colour to the brick's breaking of the window. It might be thought then, that if Davidson's theory is true, we consider the content of propositional attitudes as relevant for the purposes of explanation of action, only because it happens to co-exist with some physical

[25] See e.g. Stoutland [1980] and [1985], and Honderich [1982].

property which is responsible for the causal efficacy of events that we describe as mental. This conclusion, would certainly render Davidson's model for the explanation of actions only *pragmatically* adequate, granting it no right to claim that it captures the real explanatory relation between actions and their determinants.

The Causal Theory of Action, as construed by Davidson in his early work, does not stand or fall with Anomalous Monism, since it does not make any negative claim about the existence of laws containing psychological terms. The claim is rather that such laws *need not* exist for the causal theory to be true. However, if the criticism were correct, Davidson's causal theory would lose much of its appeal, since it would not be able to hold the middle ground between the non-causal and the causal-nomological approach any longer. To retain his thesis that reasons are causally connected to actions, Davidson would have to make some concessions to the nomological, or the functionalist theories and allow for some lawful connections between the mental and the physical. As I will attempt to show in the upcoming discussion, such a concession is radically at odds with Davidson's overall philosophical position; hence, I believe that we should view this sort of criticisms as endangering not only his Anomalous Monism, but also his theory of action, seen as a distinct causal theory.

Davidson has replied[26] to this kind of criticism by pointing out that his critics have misconstrued his ontology of events. Events do not enter into causal relations qua mental or qua physical. The mental and the physical languages are just ways of talking, ways of describing things, that make no claim about the existence of anything apart from events. Thus, Davidson rejects the distinction between causation by an event in virtue of some property **x**, and causation by the same event in virtue of some other property **y** that it possesses, as senseless. It is events 'tout court' that cause or are caused, their properties only being characteristics that we attach to them in order to describe them one way or another. So, if a statement of a particular causal interaction is true, then so will be any other statement of that interaction, provided that it uses true descriptions of the events involved; hence,

[26] See e.g. Davidson [1987] and [1993].

even the events that are (truly) described in mental terms can be causes and effects. The mental is, after all, as causally efficacious as the physical.

The Problem of Congruence in Explanation

I.

Davidson's reply is based on ontological considerations, which I will not attempt to rebut directly here, but only to undermine, by construing the arguments against him as arguments that bear on the problem of explanation, which, I think, subsists even if we accept Davidson's ontological commitments. Whether singular causation requires only the existence of events and not of their properties, is a metaphysical issue which will be decided on the basis of other considerations apart from the ones pertaining to the problem of action. This ontological thesis may be plausible, but, whether true or false, it still leaves us with the problem of explanation of action, which, I will try to show remains unsolved by Davidson's theory.

Philosophers like F. Stoutland[27] and L. Antony[28] have concentrated on Davidson's model for action explanation and have found it wanting. Of course, their criticisms take off from different standpoints and the points they make differ significantly in many respects, but, I believe, concern about the same problem underlies both criticisms. Antony calls it the Problem of Explanatory Force, whereas Stoutland calls it the Problem Of Congruence. I will use Stoutland's term which, I think, is more appropriate for the conclusion I want to reach.

Explanation contexts are non-extensional, that is, the validity of an explanation depends on the descriptions of the explanantia and the explanandum. In Davidson's theory, the onsets of desires and beliefs which are intentional descriptions of events that also have a physical description, feature in the rationalisation of the action, and thereby, are supposed to explain it. These

---

[27] Stoutland op. cit.

[28] Antony op. cit.

intentional descriptions are not used in the strict, deterministic laws that back the causal connection posited by Davidson, and hence, cannot be used as premises in a subsumptive explanation of action. It is rather the physical descriptions of these events that would stand as part of the explanans in such a covering-law explanation. This is because the physical descriptions are nomic descriptions, they classify events in types that feature in laws connecting them to other types of events. Therefore, the intentionality of desires and beliefs can play no part in a complete, Hempelean explanation of action. If Davidson's theory cannot offer an alternative, convincing model for action explanation that secures a central role to reasons, then reasons will lose their explanatory force.

Davidson's theory, makes use of two languages that run in parallel being irreducible to each other, and providing descriptions of the same entities, the concrete events. One is the language of Psychology that contains intentional terms such as belief, desire, hope, want, etc., as well as action verbs, in order to describe events that take place inside of human agents' minds and events that are classified as actions respectively. The intentional terms that are used by Psychology are not nomic terms, in the sense that they cannot, in principle, feature in strict, deterministic laws that explain and predict the events that they describe. Davidson has argued extensively for this point, but he has never denied the possibility of non-lawful generalisations expressible in the language of Psychology. However, these generalisations cannot serve as premises in covering law explanations, since, as he has argued, they contain inescapable 'ceteris paribus' clauses and hence, they could not be developed into strict, exceptionless laws.

The other language is the language of the Physical Sciences, which is considered by Davidson to comprise a closed system in the sense that it has the capacity to describe, predict and explain any event in the world,[29] using only terms from its own vocabulary. Thus, the physical language can, in principle, accommodate 'homonomic' laws, laws that is, that can be developed and refined in order to accommodate new phenomena, without resorting to descriptions outside the domain of the language. The physical language can therefore be used

[29] Or, at least, any event that enters into causal relations. See his 'Mental Events', p. 208 in EAE..

to describe the events involved in action explanations, in such a way as to subsume them under a law.

Events, in themselves are neither mental nor physical : they are just particulars that can be described either way. But, since the physical language has the capacity to describe, predict and fully explain all events, I think we are justified in saying that, in Davidson theory, it holds a privileged position. However, this privilege does not stem from the reducibility of psychological to physical predicates, because such reducibility is denied by Davidson.

II.

The fact that the two languages are detached creates a problem for Davidson's model of action explanations. Explanation is description-relative and as such, it can only be expressed in either one of the two languages. When we explain a particular action by citing the agents' reasons which, on that occasion, were also causes, we are using intentional terms. These reasons make the action seem a reasonable thing to do, by placing it into a rational pattern of behaviour. This display of rationality, does not take into account any physical considerations. For example, our evidence for ascribing intentionality to a particular piece of behaviour is not derived from knowledge of physical laws, or from observation of physical events. Our ascription depends on rational considerations, on our demand that the behaviour is understood as meaningful action, as part of a pattern that coheres with social conventions and habits, or with further actions and intentional states (beliefs, desires, expectations, etc.) that we ascribe to the agent.

All the above is, in Davidson's view, part of what action explanations consist of. But, Davidson also added the Causal Condition, in order to account for the explanatory force of rationalisations. So, in citing the reason for an action, we also denote its cause, describing it in intentional terms, in terms, that is, that do not indicate the full-blown deterministic law that covers the events in question. Davidson assumes that the full, deterministic law that subsumes the events, must do so by describing them in physical terms, and thus be a physical law. Intentional terms are then non-nomic terms, and, therefore, when we include in a singular explanation the statement that a reason was the cause of an action, we are not using any nomic property to back it. The question is then, whether such an

explanation in terms that are not nomic, can, in any sense be considered as adequate. Remember that for Davidson, an adequate explanation of **A** 's **x**-ing, should include the following statements:

i) That **A** possessed a desire for some end **y**, and the belief that **x**-ing was a means for **y**.

ii) That **A**'s desire and belief caused his **x**-ing.

The singular causal statement asserted in (ii) and the fact that (i) accords with the Rationality and Truth conditions is, according to Davidson, all we need in order to have an adequate explanation of A's x-ing. But there is another question, an epistemological one, that we need to consider here: Can we justify the assertion that the above explanation of A's x-ing is adequate? The question is distinct from the one posed above (i.e. what constitutes an adequate explanation), since it concerns the *grounds* that we have for supporting the claim for adequacy.[30] However, if no justification for this claim can be found, then the claim itself will, inevitably, become questionable; it is extremely implausible to hold that our reasons-explanations of actions are adequate, but we cannot find *any* solid ground to support any of them.

III.

Are there any grounds to support the claim for the adequacy of a singular explanation construed according to Davidson's model? It is clear that Davidson cannot justify it by appeal to strict covering laws. If I understand him correctly, his justification seems to lie in the assurance that we can be confident in our belief that a particular reason caused a particular action, without knowing their nomic descriptions. Our confidence, supposedly stems from our knowledge of non-strict generalisations (truisms) that connect the reasons with the action involved.

---

[30] Cf. D.-H. Ruben's questions of what counts as a full explanation, and of how we can justify a claim that some explanation is full, in his [1990] article. In my argument, I use the word 'adequate' to approach to what Davidson means with 'something less than explanation, that still counts as explanation'. The term is not meant in its pragmatic sense, since I don't think that Davidson holds the view that reasons explanation is only pragmatically adequate. Perhaps the only appropriate words that could be used to describe an explanation that captures the 'real' explanatory relation of the explanans to the explanandum, are 'full' and 'partial' explanation, however, in view of the pragmatic connotations of 'partial' which I want to avoid, and of Davidson's attempt to play down the importance of full explanations, I choose a deflationary term such as 'adequate'.

I believe that Davidson's arguments conflate causation with explanation. The truth of a singular causal statement may remain unaffected after substitution of a term with one which is co-referrent, but the same does not hold for singular explanation. We may be able to pick out a cause using properties that are irrelevant to the causal law that covers it, but we cannot pick out an explanans by such an irrelevant feature, because features are precisely what matters in explanation. This is stressed by Hempel, among others, when he says that 'the object of description and explanation in every branch in empirical science is always the occurrence of an event of a certain *kind*.....at a given place and time, or in a given empirical object....at a certain time'[31].

So, to justify an explanation as adequate, we have to justify our choice of features. In particular, in a causal explanation, we need to show that the description we are using is causally relevant. To pick out the cause by some irrelevant feature, is not enough for the purposes of explanation; in scientific explanation we do much better than cite the cause by using any of its true descriptions.[32] By the same token, it is a legitimate demand on any causal explanation, that it can be shown to describe the events which are related as cause and effect, in terms that are causally relevant. Presumably, one way that this can be done, is to show that the events involved, *as described*, feature in the causal law that subsumes the singular relation. Davidson, of course cannot use such strict laws to justify his causal claim. But, can we say that his claim that reasons are causes, is justified by the truisms that can be expressed in the language of Psychology?

At times, Davidson writes as if these truisms can provide the grounds for action explanations[33]. Consider the following example: We explain Tom's buying a soft-drink on a hot summer day from the kiosk, by his desire to quench his thirst and his belief that he can achieve this by buying a soft-drink from the kiosk. Davidson says, that our reasons-giving explanation is based on some non-strict

---

[31] See Hempel [1942], p. 233.

[32] Cf. also Mackie's distinction between 'explanatory' and 'productive' cause, in Mackie [1974], Ch. 10.

[33] e.g. in 'Hempel on explaining Action', in EAE, pp.261-75 and in Davidson [1987], pp. 44-45.

generalisation that is exclusively true of Tom, and which can be filled in by adding clauses that take into account all the relevant conditions that must obtain, if the action is to be carried out. So, in our example, such a truism would look like: 'If, on a hot summer day, Tom is thirsty, and he sees a kiosk, and he has the amount of money needed,....., then he will tend to buy a soft-drink from the kiosk.' Davidson stresses that such a truism would be applicable only to that particular agent, and not hold as a universal generalisation. Moreover, it has no binding force even for Tom, because no matter how many clauses we add, we can always think of some special circumstances that will not allow Tom to buy his drink, or that would even override the tendency itself.

By bringing in generalisations to support reasons-giving explanations, Davidson tries to bring his account closer to Hempel's. Of course, he still rejects Hempel's requirement that strict laws must be included in the explanans and, consequently, he also rejects Hempel's argument thesis. But can such truisms enable Davidson to justify his claim that his model provides adequate action explanations? I think not, because the generalisations in question are not sufficient to distinguish between explanation and justification, which is the problem Davidson set out to solve. The reason for this, is that the possibility of finding exceptions to the rule, may let in a case where, Tom had the reasons for buying the drink which were mentioned above, but, nevertheless, bought it for different reasons, or, did not act at all due to some set of circumstances that is not included in the clauses of the generalisation. In other words, Davidson's own argument against the non-causal theories, i.e. that they cannot account for the explanatory force of the 'because' in action explanations, would apply here too.

IV.

Hence, it cannot be claimed that truisms, as such, bring out the causal efficacy of reasons with respect to actions, since they do not help us understand the intentional properties of reasons as *causal* properties, any more than a practical inference does. It seems then that the only justification that we have for asserting the adequacy of a Davidsonian singular causal explanation, is our direct knowledge of the operation of the cause. But this knowledge cannot justify our

choice of descriptions and so, it cannot justify the causal relevance of reasons *as described.*

Consider the example given by Davidson in 'Actions, Reasons and Causes': Suppose that a hurricane, which is reported on page 1 of Monday's Times, causes a catastrophe, which is reported on page 2 of Tuesday's Tribune. A statement that asserts that 'the event reported on page 1 of Monday's Times, caused the event reported on page 2 of Tuesday's Tribune', is a true statement, regardless of the fact that it uses non-nomic descriptions to refer to the relata of the causal relation. Hence, the truth of the causal claim, does not entail that a law connecting events as described in it exists; not even the claim that 'the hurricane caused the catastrophe' entails that there is a law connecting hurricanes with catastrophes. What these claims entail, is that *some* law that covers the case exists, subsuming the related events, under *some* descriptions that are true of them. Granting Davidson these points, we can go on and ask: Can we *explain* the fact[34] that the event reported on page 1 of Monday's Times occurred, by saying that it was caused by the event reported on page 2 of Tuesday's Tribune? Obviously not.

Is the above example analogous to singular action-explanations? Perhaps an action-explanation is more like the explanation of the catastrophe in terms of the hurricane. This explanation is considerably less ridiculous than the explanation in terms of events reported in newspapers, even if, it is 'slightly less ridiculous', as Davidson says, to search for laws linking hurricanes to catastrophes. If we look at the justification for the claim that the latter explanation is adequate, we may gain some insight on how Davidson could or could not justify his model of action-explanations.

As Antony[35] points out, we can justify a singular explanation in terms of hurricanes and catastrophes, by inquiring into what kind of properties are characteristic of hurricanes and what are characteristic of catastrophes. This inquiry amounts to following an analytical explanatory strategy, as distinct from a

---

[34] I assume that the relata of explanation are not events, but facts, or statements. My assumption is in accordance with Davidson's view which is expressed, among other places, in 'Causal Relations', in EAE, p.161: 'Explanations typically relate statements, not events.'

[35] Op. cit., pp.168ff.

subsumptive one. R. Cummins[36] distinguishes between these two strategies in scientific explanation, and elaborates on the importance of each one. The subsumptive strategy is best exemplified by Hempel's model of D-N explanation. The aim here, is to explain the changes of state in a system by a causal law. These laws are provided by transition theories, theories that is, that account for the succession of states in the system.

As Cummins argues, the subsumptive strategy is by no means the only strategy which is important in scientific explanation. What is often required, is an analysis of the properties of a given system and this requires what he calls a property theory. The analytical strategy, thus seeks to explain what is it for a given system to have a certain property. Of course, this is important for explaining state transition also, since it is through property theories that we can explain in virtue of what attributes do pairs of events stand in the relation of cause and effect.

A most interesting, and relevant to our discussion, application of property theories, is to explain dispositions. The question, 'what is it for object **o**, to have disposition **d**?' is answered by way of analysing how **d** is instantiated in **o**. Such an explanation, requires property theories to provide 'instantiation laws' which state the necessary individual 'components' that any object must have, in order to instatiante the property **d**. An instantiation law will therefore tell us what conditions must a certain system fulfil, in order to manifest the disposition.

We can similarly think of hurricanes as having the dispotitional property of producing disasters. But our confidence in the *hurricane's* relevance in a *causal explanation* of the *disaster*, is not based on knowledge of the lawlike regularity 'hurricanes cause disasters'. Our confidence is based on our (perhaps rough and incomplete) knowledge of the way that the disaster-producing property is instantiated in hurricanes. We know, for example, that we can inquire into what sort of events make up a hurricane, and what sort of consequences would a collection of such events have. On the other hand, a similar analysis of disasters would bring up their relevant properties and reveal how can events with such properties be caused by the collection of events making up a hurricane. Of course, this would require the help of transition theories, i.e. theories that provide causal

[36] In Cummins [1983], Ch. I.

laws linking events at the more 'fundamental' level of description. But, by no means would the transition theories suffice to back our claim that the occurrence of the hurricane causally explains the occurrence of the disaster. Property and transition theories have a complementary role in the justification of our singular causal explanation.

Therefore, we can see that our confidence in the claim that the occurrence of the hurricane causally explains the occurrence of the disaster, does not lie on our knowledge of the singular causal statement. To have grounds for accepting such an explanation as adequate, is to know that some property theory can be formulated, a theory that will explain *why the things that have* the dispositional property of being a hurricane *have these kind of effects*, whereas, *other things do not.* The concepts of hurricane and disaster allow confidence that such a fruitful inquiry can be made. The concept of an event reported on page 1 of the Times, does not.

V.

It must be clear, by now, why Davidsonian action explanations cannot be analogous to explanations of disasters in terms of hurricanes. Davidson denies that there can be any sense in the question 'in virtue of what properties does event c cause event e?' But without the qualification 'in virtue of', how can we make sense of a reason's explaining an action? Is the mere fact that we affirm the causal relation enough to give us confidence in a singular causal explanation of action in terms of its reasons? That this is not so, is I think manifested in the problem of deviant causal chains. There, Davidson's analysis fails to distinguish between deviant and non-deviant causes of behaviour, precisely because it cannot account for a desire and belief's being causes in virtue of their 'rationalising' characteristics and not in virtue of their 'disturbing' ones. Hence, in a singular action explanation, merely citing the 'producing cause' is not enough. We need to be able to justify our choice of descriptions by appeal to some property theory, if no causal law can be formulated in their terms.

Could Davidson's theory allow for the development of such a property theory? The first thing to note is that, as Cummins points out[37], the claim that a property

[37] ibid. pp.22-6

theory identifies physical instantiations of psychological properties, is distinct from the claim that it carries out reductions. A certain psychological property, can be instantiated in a variety of physical systems (all of them sharing a given structure in virtue of which they manifest the property), which have disparate physical properties; in that case, no reduction of the psychological to a physical type will be possible, except if we accept reduction to open disjunctions of physical types as genuine reduction. However, the instantiation laws that analyse systems possessing a given property, into individual components with a certain mode of organisation, do not describe mere correlations between properties and structures. They are *laws* and they are themselves derivable from nomic attributions, i.e. lawlike statements which attribute properties to the individual components of the organised system, in virtue of which the higher-level property is manifested in that particular system.

The above exposition of Cummins' analysis of the explanatory role of property theories, is meant to put across the point that, in order to show that a certain property is causally relevant, it is not necessary to reduce it to one that features in a causal law. (Which is something that Davidson's theory certainly does not allow) On the other hand, for an analysis of a given property in terms of its instantiations to be successful in bringing out the property's causal relevance, more than mere statistical correlations are needed; in particular, an adequate theory must show the property's relevance, in the light of *nomic* attributions and *laws* of instantiation that apply in the particular case.

So, given that property theories do not necessarily license reductions, could we say that Davidson's theory might be developed along the lines of a functionalist position, as sketched out above, so that the causal relevance of intentional attitudes is justified? I believe that his arguments for the anomalism of the mental, insulate psychological terms (and hence action descriptions), not only from outright reduction to physical ones, but also from any lawful connection of the kind that a property theory would require if it was to have any explanatory value at all. In 'Mental Events', he writes: '[T]here may be *true* general statements relating the mental and the physical, statements that have the logical form of a law; but they are not *lawlike*'.[38] Thus, it would be impossible, according to Davidson, to find any

[38] 'M.E.' in 'EAE', p.216

genuinely lawful statement that would specify in virtue of what physical nomic attributes of the components of a certain type of structure, a given psychological property were instantiated.

Davidson's rejection of the possibility of formulating genuine psychological or psychophysical laws stems from his conviction that the ascription of intentional attitudes and behaviour to agents, on which psychological explanations are based, is a process which is necessarily holistic, interpretative and normative. Its holism derives from the fact that we cannot ascribe beliefs, desires, expectations, etc., to agents, without taking into account their other past, present and future attitudes and behaviour. We have to make our ascriptions cohere with the pattern of the agent's other attitudes.

But holism alone is not sufficient to differentiate Psychology from the Physical Sciences, since it is their feature too. What makes the difference is the interpretative and normative nature of the considerations on which these ascriptions are based. When we try to determine what a person's attitudes and aims in behaviour are, we have to *interpret* their meaning, so that the holistic requirement is satisfied: 'It is not merely, as with the measurement of length, that each case [of assignment of intentional attitudes] tests a theory and depends upon it, but that the *content* of a propositional attitude derives from its place in the pattern'.[39] So, the special feature of psychological ascriptions is not merely that they are interdependent, but that they derive their content from the attitudes on which they depend. We interpret other people's speech and behaviour so that they make sense against a background pattern of behaviour, which is verbal or otherwise. Our interpretation, moreover, has a normative character: We fix people's attitudes, by constraining what they *ought* to believe, desire, etc. given this background pattern.

The interpretative, holistic and normative character of our inquiry into people's attitudes and behaviour, derives from Psychology's commitment to the 'constitutive ideal of rationality'. Although agents may be more or less rational, we must interpret them according to the set 'ideal' of rationality, basing our interpretation on what they ought to believe, desire etc., given our ascriptions of

[39] ibid., p.221 (Emphasis and parenthesis mine).

background attitudes. Otherwise, we can make no sense of their actions and self-avowed attitudes. This commitment to the 'constitutive ideal of rationality' is forced upon us if we are to treat men as rational agents and thinkers: 'If we are intelligibly to attribute attitudes and beliefs, or usefully to describe motions as behaviour, then we are committed to finding, in the pattern of behaviour, belief and desire, a large degree of rationality and consistency.'[40]

So, what excludes the possibility of formulating lawful connections between the physical and the mental, is Psychology's commitment to the 'ideal of rationality' which is an essential part of our concept of intention, belief, desire and action and cannot be compromised in the face of empirical evidence. The decision to treat men as 'rational agents with goals and purposes and as subject to moral evaluation'[41] , is a choice that does not allow us to let Psychology become a 'hostage to empirical fortune'. Hence, we cannot make an arbitrary choice of our scheme of interpretation (as we do with physical theories which are underdetermined by data), and establish lawlike connections between the mental and the physical, because that would leave the 'ideal of rationality' open to falsification by physical evidence: A law connecting mental predicates with physical ones, would have to be subject to empirical falsification, thus rendering psychological ascriptions of intentional attitudes sensitive to physical evidence. Hence, the 'constitutive ideal of rationality' would lose its *a priori* status, which Davidson considers as an indispensable part of our commitment to 'viewing men as rational agents'.

'The constitutive force in the realm of behaviour derives from the need to view others, nearly enough, as like ourselves. As long as it is behaviour and not something else we want to explain and describe, we must warp the evidence to fit this frame. Physical concepts have different constitutive elements. Standing ready, as we must, to adjust psychological terms to one set of standards, and physical terms to another, we cannot insist on a sharp law-like connection between them'[42].

---

[40] 'Psychology as Philosophy', EAE, p.237

[41] ibid., p.239

[42] Ibid. p.239.

It is not my purpose here to criticise Davidson's arguments for the anomalism of the mental; my point is rather to show that, as Antony argues[43], his scepticism regarding psychophysical laws does not concern only type-type reductionism, but any attempt to formulate nomic connections between the mental and the physical, connections that might make ascriptions of intentional attitudes subject to falsification by physical evidence. Evidence which is pertinent to physical laws and attributions, is not to be admitted to play *any* role in the process of interpretation of the agent's behaviour. If it were admitted, that 'would amount to changing the subject',[44] i.e. using non-mental descriptions to explain what we now characterise as intentional action.

It is now evident, I believe, that Davidson would not accept the possibility of lawfully connecting psychological properties to physical systems, as instantiation laws would require: 'mental and physical predicates are not made for each other'.[45] A property theory that would explain how a pair of desire and belief cause, *in virtue of their intentional properties*, a piece of behaviour, would have to match up in a law-like statement the property of having a desire to.....and the property of having a belief that....., with a certain mode of organisation of the physical components of a system. Even if this instantiation law would not necessarily license a type-type reduction, it would nevertheless provide sufficient conditions, in physical terms, for these intentional properties to be instantiated. But that would compromise the (as proclaimed by Davidson) 'a priori' character of our ascriptions of intentionality, because it would make them liable to falsification by physical evidence: we would be in the position to tell, by examining the physical state of the brain, whether a certain belief or desire was instantiated, and hence it would be possible to overrule ascriptions based on rational considerations. This is something that, according to Davidson, cannot be allowed if we are to describe human behaviour in intentional terms, and thus view men as rational agents.

Therefore, I think that Davidson's unified Theory of Mind and Action, does not have the resources to admit further explanation (and thus justification) of the causal relevance of reasons with respect to actions. His claim that singular action

---

[43] op.cit.174-83.

[44] M.E., p.216.

explanations are causal explanations, loses its plausibility in the face of his theory's inability to provide grounds for justifying it. According to this theory, desires and beliefs are causes of actions, but it is impossible to ever come to know why did a particular reason cause a particular action. We may know that it did (since we may directly know that singular causation is at work even if we do not know the relevant description), but no further explanation could be possibly admitted. But this surely is not enough to ground Davidson's causal condition. As Cummins' discussion of scientific explanation will have made clear, in giving scientific explanations, we can normally do much better than cite the 'producing cause' of an event. The concept, and the actual practice of explanation, is much richer than Davidson allows it to be in the case of action.

VI.

Davidson proposed his Causal Theory of Action, as a theory that made sense of the difference between merely having a reason *and* acting, and acting *because* of it. The Causal Condition was supposed to account for the explanatory force of the 'because' in reasons-explanations. But, as we have seen, the Causal Condition cannot confer additional explanatory force upon the reasons, since the assertion that the onset of a reason causally explains the action, cannot be grounded on any causal law, or any intentional property of the reason shown to be causal by an appropriate theory.

Hence, it must remain a mystery, in Davidson's theory, why a particular reason caused a particular action. The link between reason and action seems to be congruent with the link between physical events that take place in the brain, and 'pure physical movements' of the body. It is prima facie natural to believe that, every time that an agent performs an action for a reason, there are physical events that start in her brain and end, through a chain of causal connections, in movement of some part of her body. The central problem in the Philosophy of Action, is to give an explanation of action that helps us understand how this congruence between rationalisation and event causation is possible. To combine, that is, in a harmonious and convincing way, the rational order of human behaviour that is

---

[45] ibid. p.218.

described and explained by the 'Human Sciences', with the natural order that is being charted with such success by the Physical Sciences. Davidson's theory fails to account for the problem of congruence, because it does not accept any kind of systematic correlation between mental and physical types. So, the fact that a particular reason caused a particular action is, in his theory, unexplainable. It is a brute fact: There is no connection to be found between the intentional kind that rationalises the action, and the physical kind that causally explains it.

Davidson might reply that the problem of congruence is only a pseudo-problem for him, since his theory is a monistic theory that takes the onsets of reasons to be *identical* to the physical events that are revealed as causes by physical science. Then, there is no congruence between the two connections (reasons-actions and neurochemical events-movement), because there are no two connections, but only one, involving two entities, and not four. Then, no explanation of why a particular reason caused a particular action is needed, because the reason *is* the physical event that features in the relevant causal law.

But this response does not take into account some of the real concerns that we have when we seek an action explanation. When we explain an event, we explain it under a description, and when two descriptions are equally true of it, it is not senseless to ask for an explanation of why this is so. This is attested to by any theory that does justice to the epistemic as well as metaphysical considerations pertaining to explanation. Davidson sometimes writes as if we can explain the 'pure occurrence' of an event, without involving any of its characteristics in the explanation, i.e. as if we can explain why **b**, which *happens to be* **F**, occurred, without explaining why an *F event* occurred. Such an explanation, if it makes sense at all, does not satisfy our original demand, which was for an account of a specific characteristic of the event in question, in the case of action, of its intentionality. If a Davidsonian action explanation is understood as an explanation of the 'pure occurrence' of an event, that *happened to be* an action, then the theory has failed its original purpose. The theory has to provide an explanation of why the physical movement examined has a certain property, namely *of being a so and so action*. It should also provide us with a general account, of how it is possible for a given event, to have at the same time the property of *being a so and so action*, rationalised by a certain reason, and of *being a certain physical movement*, caused

by a given physical event (the problem of congruence). A 'pure occurrence' explanation can do neither of the above, and so it fails as an action explanation.

If, on the other hand, Davidson's theory is supposed to provide 'property explanations', that is, explanations of the type 'why **b** is **F**', then it fails on the grounds that I provided earlier that the model proposed by the theory, cannot be justified as adequate: Reasons are unsupported as causes and hence, the necessity of the Causal Condition remains undemonstrated.

Finally, T. Nickles[46], has observed that events that are described in mental terms, are not explainable on Davidson's theory, even if their particular identities with physical events, and the relevant strict causal laws are known. The reason is that, "no particular identity can take us, in an explanation-preserving way, from '**Fb**' to '**Gb**', or from '**(∃!x) Mbx**' to '**(∃!x) Sbx**', not even the identity '**(ιx) Mbx = (ιx) Sbx.**'" Explaining why **b** is **F**, does not explain why **c** is **G**, even if we know that **b=c**. I think that this observation touches the heart of the problem that Davidson's theory of action-explanation faces: It cannot give an explanatory account of *properties*, as opposed to *'pure occurrences'* of events, because, having each foot on a different level of support, it cannot provide a unified, solid justification for the adequacy of its explanations.

Conclusion

It maybe the case that, on other grounds, Davidson's Causal Theory of Action and his Anomalous Monism, are deemed metaphysically plausible, but the problem that they face as theories of explanation, (in the limited domains of action and psychological explanation), suggests, I think, that they are infected at the metaphysical level too. Perhaps their ontology is too austere to capture the richness of ordinary psychological explanations, or their analysis of intentional concepts is simply off the mark. There have been many suggestions as to what is

[46] See Nickles [1977].

the problem in Davidson's metaphysics. F. Stoutland argues[47] that the source of the problem in the Causal Theory of Action, is its identification of action with physical movement. We ascribe intentionality to bodily movements, says Stoutland, only insofar as we understand them as action. And to do this, we have to see the action in the light of reasons. Thus, there is some conceptual dependence between action and its reasons, that does not allow us to view them as separate entities, bearing the relation of cause and effect to each other. Hence, one cannot enquire about necessary conditions for a given physical movement being an action, since, we cannot even isolate a particular piece of behaviour without first understanding it as action. Hence, the *understanding* of action is prior to its *explanation* by reasons, and a theory that does not respect this priority, is bound to lead us into confusion. Von Wright's theory, which I am going to consider in the second part, offers such an alternative understanding of action.

[47] In Stoutland [1985].

PART II : THE INTENTIONALIST APPROACH

CHAPTER 4 : VON WRIGHT'S THEORY OF ACTION

Von Wright's account of intentional action, expounded mainly in his book *Explanation and Understanding*[48] and further articulated and defended in a number of subsequent publications[49] does not amount to an *analysis* of action, as Davidson's theory does. His theory is rather an attempt to set out the nature of the relations than hold between the concepts of intention, belief and intentional action, by discussing a theoretical schema, the Practical Inference schema that, according to him, provides the basis for the explanation, understanding and prediction of human action. In all that, the concepts of intention and intentional action are taken to be just as basic as that of event-causation[50], and hence unanalysable in terms of it.

In this part of my thesis, I will try to present and assess the conceptual framework of intentional action as it is developed by von Wright, by first setting out the terms and the explanation schema used by the theory, and then by examining the nature of the conceptual connection between intention and action that is claimed, in von Wright's theory, to be the basic explanatory relation. Finally, I will follow two different lines of criticism against von Wright's intentionalist approach and assess their impact on it.

Preliminaries

Unlike the causal theory, von Wright's model for action-explanation does not analyse action in terms of causation between events. The action that is standing

[48] Von Wright [1971], (henceforth E&U).

[49] See for example von Wright [1972], [1974], [1976], [1980].

[50] In fact, von Wright understands the concept of event-causation as being dependent on that of action (see e.g. E&U Ch. II and von Wright[1974]), but his views on causation will not concern me here.

in the explanandum sentence, cannot be identified with an event with a special type of causal theory. It is rather a piece of behaviour, which is already intentionalistically understood and which cannot be accounted for by using any other concepts.

However, the concept of action has some internal structure. The most important distinction in von Wright's account, is that between the *result* and the *consequence* of an action. The result is that state of affairs or event which the agent intends to bring about by his behaviour.[51] It is logically intrinsic to the action, since, in the case that it is not brought about, we cannot say that the agent has successfully performed the action under the description featured in his intention. What we can at most say, is that the agent tried but failed to perform it. Consider for example an action of mine, say throwing a brick towards a window. If my intention in throwing the brick towards the window is to break it, then the result of my action is the breaking of the window. If I fail to hit it, then I cannot be said to have performed an action of breaking a window, but a different one, e.g. that of throwing of a brick or of *trying* to break a window.

Most actions will also involve events that are either causes or effects of the result of the action, but which do not feature directly in the agent's intention. Their occurrence is not a necessary condition for the successful performance of an action, because they are not *intrinsic* to it. The agent may or may not know about these *causal antecedents* and *consequences* of the result of his action and may or may not want them to materialise. This knowledge is of course very important in the ascription of moral responsibility, but is irrelevant to the present context. In our brick-throwing example, a consequence of my intentionally breaking a window may be that a certain person sleeping inside wakes up, or that a certain glass-maker gets richer. A causal antecedent of the result of my action may be the projectile of the brick, or my arm's movement.

---

[51] The result of an action does not need to be a change, since von Wright allows for preventive actions, where the agent prevents a change from taking place, and forbearances, where the agent forbears to do something which was in his power to do. In my discussion of von Wright's theory, I will concentrate on productive actions, where the agent brings about some change through the movement of his body.

Note that the distinction between the result and consequence is relative to the agent's intention: If all I intend by my behaviour is to throw a brick in the air, then the result of my act will be the projectile of the brick and the window breaking will be just a consequence. On the other hand, if my intention is to wake up the man sleeping inside the house, then his waking up will be the result of my act, and the window breaking and the brick throwing will be its causal antecedents. Using Davidson's terms, we may say that by giving the result of an action, we are implying a description under which the action was intentional. However, it is important not to be mislead by this terminology. For von Wright, actions are not events that take on different descriptions. An action, is rather an instance of an agent's bringing about an event[52]. Hence actions are to be identified with the *instances of a relation* that holds between agents and events, and not simply with the latter. So, when we speak of different descriptions of an action in the context of von Wright's theory, we are really speaking of different actions.[53]

Von Wright also distinguishes between *doing* something *intentionally* and *intending to do* something. Whenever we do something that we intend to do, we do it intentionally. But it is not true that everything that we do intentionally, we intend or intended to do it. For example, when I intend to make myself a cup of coffee, I put water in the kettle, place it on the fire, remove it from the fire etc., and I do all these things intentionally, having the preparation of a cup of coffee as the object of my intention. But we cannot say that I *intended to do* all the things that I *did intentionally*. I may not even have had complete awareness of my actions.

According to von Wright, the concept of intentional action is more basic than that of non-intentional action. When we do things non-intentionally because of a mistake, or because of lack of knowledge, or by negligence, there is also an intentional action that we have performed. If, for example, I unintentionally wake up a person sleeping in his house, this is something I did by performing an intentional action, e.g. breaking his window with a brick. This priority given to

---

[52] See von Wright [1963], pp.35-6.

[53] See his reply to Donagan, in von Wright [1984], p. 810.

intentional action is in accordance with Davidson's view who, as we have seen, holds that non-intentional actions are just intentional ones under different descriptions. Although he accepts that non-intentional action is 'parasitic' upon intentional action, von Wright still allows for the existence of the former, in cases like laughing, yawning, etc. However, even these can be called actions only because we can learn how to do or omit to do *intentionally*[54]. If this were not the case, as happens for example with breathing, then we would not call them actions at all. Hence, all actions either involve something done intentionally by the agent, or something which *can* be performed or omitted intentionally[55].

There is one more distinction that should be made, and this concerns the term behaviour. The term sometimes is used ambiguously by von Wright . It may be meant as 'intentionalistically understood' or 'action like' behaviour, in which case it can feature in the explanandum sentence of a teleological explanation, because it is already understood as action and not as mere bodily movement. This last sense, i.e. behaviour as mere *bodily* behaviour (or movements), is what is meant by causal theorists when they analyse action in terms of a certain piece of behaviour caused by a certain type of mental event. Von Wright too sometimes means it in this sense. On my part, I will use 'behaviour' in a neutral, non-committal sense that may mean *either* 'action like' behaviour, *or* 'mere' behaviour. If I want to talk of behaviour exclusively in one of these two senses, I will explicitly talk of 'action-like' or 'mere' behaviour.

Finally, in my discussion of von Wright's theory I will avoid, as I did with Davidson, entering into the problem of the individuation of action. For every would-be 'action designator', i.e. for every sentence purporting to designate an action, we can examine whether it really describes a piece of intentional acting on the part of the agent, leaving aside the question of how many actions are designated by act-designators that are somehow connected. In other words, for

---

[54] See von Wright [1980], p.18

[55] In view of the priority given by von Wright to intentional, over non-intentional action, I will, from now on, use the term action to refer only to intentional action. If I want to talk about non-intentional action, I will make it explicit.

every action-description given, we can discuss its intentionality without making any particular claims as to whether the descriptions we have are descriptions of a single act, or they are descriptions of distinct acts.

The Practical Inference Schema

After considering various versions of the third-person Practical Inference Schema (hereafter P.I.-schema), which he finds to be inadequate, von Wright arrives at the final formulation:

'From now on A intends to bring about p at time t.

From now on A considers that, unless he does q no later than at time t′, he cannot bring about p at time t.

Therefore, no later than when he thinks time t′ has arrived, A sets himself to do q, unless he forgets about the time or is prevented.'[56]

The first thing to note here is that (as von Wright himself has pointed out[57]), the term 'Practical Inference' is a misnomer. With his P.I.- schema, he does not purport to provide an account of practical reasoning, as for example E. Anscombe has done. His aim is rather to unravel the structure of the concept of intentional action, by articulating a theoretical schema which he considers to play the central role in the explanation, understanding and prediction of human action. But he cannot, and does not claim that the Practical Inference presented here gives us a description of the actual process of deliberation of agents. So, many criticisms of the schema on the grounds that it does not depict accurately the process of rational deliberation[58], are beside the point, since this is *not* what the schema is supposed to do.

---

[56] See E&U, p.107.

[57] See his reply to Anscombe, in von Wright [1984], p.821

[58] For a criticism of this sort, with many interesting insights on practical reasoning, see Anscombe [1984].

Von Wright's aim is rather to set out an 'ideal' standard for the explanation, understanding and prediction of action, without discarding or reducing to something else the intentional concepts that are used in ordinary psychological descriptions and explanations. He regards the P.I. scheme as providing a model of complete, or 'ideal' explanation in the 'Human' Sciences, in roughly the same way as the Deductive-Nomological model serves as a paradigm of complete explanation in the 'Natural' Sciences. He writes:

'Broadly speaking, what the subsumption-theoretic model is to causal explanation and explanation in the natural sciences, the practical syllogism is to teleological explanation in history and the social sciences'[59].

The P.I.- schema, if 'turned upside down', becomes a schema for the teleological explanation of action, i.e. the conclusion becomes the explanandum and the premises become the explanantia. So, the fact that agent A did (or set herself to do) q, is explained in terms of her purpose to bring about p and her belief that q was the means to achieve this. The teleological explanation that emerges is intentionalistic in character, in the sense that it makes essential reference to the agent's intentional attitudes. Hence, what matters in such an explanation, is not what the objective state of affairs is, but how it is *seen* to be by the agent. In that respect, it differs from ordinary teleological explanations that may be applicable to non-intentional systems.[60].

A second important use of the P.I. schema is for the understanding of action. Since the explanandum of a teleological explanation is a piece of behaviour 'intentionalistically understood', it follows that the explanation of behaviour presupposes that we have first understood behaviour as action. In most normal cases, 'we say off-hand of the way we see people behave that they

---

[59] E&U, p.27

[60] This point was made by Ch. Taylor in his book *The Explanation of Behaviour*: 'We can thus see the full extent of the difference between explanation by purpose and the type of teleological explanation which would apply to our imaginary physical system. For in the former case the teleological account holds not of the organism in its «geographical» environment, but of the agent in his «intentional environment», the environment as it is for him. Thus the notion of a centre of responsibility is integral to our account.' (Ch. Taylor[1964], p.62.)

perform such and such actions'[61]. These are actions that are familiar to us from our personal experience or from having observed other people perform them on past occasions. For these every-day actions, one does not normally need to interpret them by using a P.I. Von Wright's contention is that, in these cases, we do not need to place an interpretation on something 'neutral' that we see: We are already 'perceiving' it as action. Only when we look for an explanation, i.e. for a further purpose in the action do we attempt to construct a P.I. But there are cases where we may doubt whether the behaviour that we see is intentional or not. Consider the example given by Stoutland:

' We see an agent making some movements at a window. Is he doing anything intentionally? If so, what? At this point we cannot tell by seeing... Then we see he has a glass cutter. A practical inference emerges:

He intends to get into the house.

He believes that he cannot get into the house unless he cuts a hole in the window.

Therefore, he cuts a hole in the window'[62].

In the example given, a P.I. guides our understanding the behaviour as action. We could not say off-hand whether he was doing anything intentionally or not. What we needed was to advance some plausible hypothesis as to what the possible aim in the behaviour might be. From the fact that a possible purpose in his behaviour emerged when we saw the glass cutter, we were able to *understand* the behaviour as intentional through the construction of a P.I. that could possibly explain his action. After having constructed it, we were able to 'see' his movements as aiming at certain result, e.g. to cut a hole in the window without being seen. So the understanding of behaviour as action is the second use in which the P.I.-schema may be put.

The third use for the schema is that of the *prediction* of action. If we stand 'ex ante actu', we can use the P.I.-schema in order to predict what the agent's behaviour will be, provided that we know what her intentions and beliefs are. If we know that she intends to bring about p and that she thinks that unless she

[61] von Wright [1972], p.32.

does q, she will not achieve this, we can predict with varying degrees of reliability (depending on the particular circumstances and the reliability of the agent), that she will do q. The character of this kind of predictions will be discussed in the next chapters, where the nature of the connection between the premises and the conclusion of the P.I. is examined.

After considering various examples where the agent did not act according to her pre-formed intention and belief, von Wright arrived at the final formulation of the schema presented above, which he thought comes as close as possible to being logically conclusive. He introduced the parameter of time, in order to ensure that the agent keeps and remembers her intention up to the moment of action (if the inference is to apply to her), and added the 'unless prevented' clause, in order to exclude cases where factors beyond the agent's control prevent her from carrying out (or initiating) her action. Now, von Wright seems to think that these qualifications are sufficient to ensure that the P.I. is applied only on 'closed systems', on systems, that is, where there is no outside interference that affects the relation between intention, belief and action. Whether he is right or not is a matter of dispute but it will not be of relevance here.

There are writers[63] that have used the locution 'normal conditions' to group together all the specific circumstances that must hold for the P.I. to apply, e.g. the agent's being able to do q, her knowing how to do it, the particular facts of the situation allowing her to do it, etc. I will not discuss what else should, or could be included under the heading 'normal conditions', taking for granted that some formulation where all the relevant factors are accounted for, is possible. In any case, I think that most of the clauses suggested by other writers, are already taken into account by von Wright, if only implicitly. For example, the agent's belief that she knows how, and is able to, do the action q, is implied by her intention to do p and her belief that doing q is necessary for bringing about p[64].

---

[62] Stoutland [1984], p.313.

[63] See e.g. Martin [1976],p.328. Also Churchland [1970], discusses the P.I. schema, (treating it as a law) and points to the considerations that might lead us to suspend application of the schema on a particular occasion.

[64] See e.g. E&U, pp.100-3.

Also the condition that the specific facts and circumstances of the situation must allow the agent to do q is, I think, implicit in the clause 'unless prevented'. In any case, I will assume that the 'normal conditions' that must hold for the inference to be valid, can be filled in by some ideal theory.

Finally, a common criticism made against von Wright's final formulation of the schema, is that it fails to take into account intentional explanations in terms of sufficient, rather than necessary means for a given end[65]. Von Wright thinks that, for the inference to be conclusive and for the teleogical explanation (which is just the 'converse' of a P.I.) to be complete, we must insist that the action q that the agent performed in order to bring about p, was thought by her necessary for p. Von Wright allows for cases where agents act on sufficient reasons but he thinks that these actions can be either explained as necessary for some further intention, or, if there is no such intention, their explanation must stay in a sense incomplete. If for example A thinks that she can bring about p by doing either q or s (either one of them is thought sufficient for p), and she does q say, then we can say either one of two things about her action. First, we may say that there was some criterion of choice for the agent (e.g. that q would save her time, effort, etc.) and hence, that her choice can be explained teleologically in terms of some further intention (e.g. to bring about p with the least effort possible). The explanation will now be a complete teleological explanation, based on a conclusive P.I.. Otherwise, if there was no consideration whatsoever favouring the choice of q over s, then we must say that the explanation of her doing q in terms of her intention to bring about p, is incomplete. Then, the only conclusion that will logically follow from the premises, will be that 'the agent does q or s', and no further explanation for her choice to do q will be possible. Von Wright does not deny that agents sometimes act in this manner, i.e. choose between competing means for a given end in a completely fortuitous way. He just thinks that these actions cannot be *completely* explained.

[65] See e.g. Anscombe [1984] and von Wright's reply, in von Wright [1984] pp.819-824.

CHAPTER 5 : THE LOGICAL CONNECTION ARGUMENT

A First Version of the Logical Connection Argument

One of the most disputed issues in the Philosophy of Action, concerns the nature of the connection between the premises and the conclusion of a Practical Inference. Most 'non-causalists' have developed some version of the Logical Connection Argument, in order to prove that there is a logical relation between the volitional and cognitive attitudes that feature in the premises of the P.I. on the one hand, and the action that features in its conclusion on the other, and hence show that the mental attitudes mentioned cannot be causes of the action.

The Logical Connection Argument (hereafter L.C.A.), in its most general form, is based on two premises. The first premise is taken to be an uncontroversial element of the Humean view on causation, namely that causes are contingently related to their effects. The second premise asserts something which is taken to be part of our concept of intentional action: That the reasons which rationalise an action, are logically (non-contingently) related to it. The conclusion drawn by non-causalists, is that the causal theorists must hold inconsistently that the reason which explains an action is both contingently (i.e. causally) and non-contingently (i.e. logically) related to it. Therefore, the argument concludes, the Causal Theory must be wrong.[66]

The argument, in its original formulations[67] attempts to establish that between an intention (or act of volition)[68] and its object, holds a type of logical

---

[66] Each one of the two premises and the conclusion, as mentioned above, is in need of some qualifications, which I will make for each particular version of the argument that I will examine.

[67] By e.g. Melden in his 'Free Action'. See Melden[1959].

[68] Although Davidson initially did not think that intentions could be causes of actions, in view of the challenge presented by incontinent actions, he later came to recognise that they should be. (See his 'How is Weakness of the Will Possible? and 'Intending', in EAE and Davidson [1987]. See also Bishop [1989], Ch. 3 for an argument to the effect that Davidson *needs* to accept them as

connection which Stoutland calls 'weak' type[69]. The weak type of logical connection between the intention and its object, consists in the fact that the only way there is to 'see' an intention, is to 'see' its object. In other words, the only way one can intrinsically describe and identify an intention, is by saying what it is an intention of.

The argument in this form has been shown to be inadequate by several philosophers.[70] Stoutland argues convincingly that Melden's defence of the L.C.A. does not go beyond establishing a weak type of logical connection between intention and its object, and hence that it fails to prove the untenability of the causal theorist's position. It is true that a plausible causal theory must accept that the intention has no intrinsic feature apart from its content, and therefore, that it is logically related to its object in the weak sense: A theory of intention that characterises intention in terms of some 'inner' impression left on the agent, and recognised by him through introspection, is logically possible, but extremely implausible, and thus it would not provide a sound basis for the establishment of a viable Causal Theory of Action. Therefore, the causal theorist must agree that intentions bear the weak type of logical connection to their objects. But it does not follow from this that the causal relation is excluded. It can still be a contingent fact that my 'intention to q' results in q, even if the only way we can characterise it is by referring to its object. The existence of the weak type of logical relation is compatible with the *occurrence* of the intention being a (contingent) cause of the *occurrence* of its fulfilment, i.e. its object. So, the existence of the weak type of logical relation, does not rule out the possibility that a causal relation also connects the *events* corresponding to the occurrence of the intention and its object. Since the proponents of this early version of the L.C.A. did not go beyond the establishment of the weak connection, their argument does not have any force against the Causal Theory of Action.

---

causes). In view of this, recognised by Davidson, need, I will construe the LCA as an argument against the possibility of intentions, rather than desires, being causes of actions.

[69] See Stoutland [1970], pp.120-2

[70] Most influential has been the rebuttal by Davidson in ARC..

Von Wright's Verifiability Argument

I.

Von Wright agrees that the establishment of the weak type of logical connection is compatible with the occurrence of the intention being a cause of the occurrence of the act[71]; yet, he still thinks that there is some sort of logical dependence between intention and action, a dependence that rules out the possibility of a theory which construes intentions as Humean causes of action. This 'non-causal' intuition is expressed much more convincingly through a complete theory of action such as von Wright's. The L.C.A. argued from the standpoint of a complete theory can be brought to bear on the causal theory more forcefully than any purely 'negative' argument. So, in his version of the L.C.A., the practical syllogism which is the basic schema for the explanation of action, plays the central role.

The way he tries to establish this logical dependence, is by means of the 'verifiability argument'.[72] This is an argument that has been criticised for being either circular or inconclusive[73] by some of his critics; even von Wright's own position regarding its conclusiveness has wavered on a number of occasions. Prima facie, it seems that von Wright attempts to prove that there is a logical dependence between the singular propositions that serve as premises in the P.I.-schema (taken as a conjunction), and the proposition that serves as conclusion. To do that, he has to show that at least one combination of truth-value assignments to the propositions, is not logically possible. That is, he takes as a sufficient condition for the logical dependence, that at least one of the conjunctions p&q, ~p&q, ~p&~q, p&~q is a contradiction (where p is the conjunction of the propositions that serve as premises). He tries to show that this condition is satisfied in the P.I., by arguing that the premises and the conclusion cannot be verified independently from each other, and using the additional assumption that for two propositions to be logically independent, it

[71] See E&U, p.94.

[72] The argument is found in E&U, p.94ff.

must be possible to verify or falsify each one independently of the other. It should be stressed here that von Wright is talking about the *logical* possibility of verifying or falsifying the propositions in question, and not about the *factual* or *pragmatic* possibility that depends on our epistemic capacities[74].

II.

The first part of the argument considers the problem of the verification of the conclusion of a given P.I.: How can we establish that the conclusion 'A does q'[75], of a practical syllogism which serves as the basis of a provisional explanation[76], is true? Von Wright points out that, in order to do this, it will not be enough to establish that the result of the action attributed to A has materialised, even if we can assert it came about as an effect of A's bodily movements. These movements, will constitute an *action* on the part of A, only if there is a further aim which the agent wants to achieve by performing them. Hence, to verify that 'A does a', we need to consider various purposes and beliefs that the agent may have, which are in accordance with the behavioural manifestations that we observe. In other words, to verify the conclusion of a

---

[73] See e.g. Tuomela [1977] pp.185-91 and Martin [1976].

[74]Still, even if one sees the verifiability or falsifiability of the propositions in the above sense, von Wright's assumption is questionable. As Tuomela (op.cit. p.185) observes, two propositions in the context of a scientific theory may not be independently verifiable, but still have the required independence. I will set aside this problem and go along with von Wright's assumption, since, in any case, as I will try to argue subsequently, his argument does not *prove* the mutual dependence of verifiability, and therefore the assumption is not, after all, crucial.

[75] For the purpose of simplicity I will ignore the complication presented by the inclusion of the time factor and the 'unless prevented' clause in the conclusion, and assume that, with respect to them, 'normal conditions' prevail. This simplification will not affect the exposition of von Wright's main point.

[76] In von Wright's own discussion, it is not clear whether he is talking about the verification of the conclusion of some P.I. which is offered as a provisional explanation, or whether his argument concerns the verification of an ascription of intentionality to a piece of behaviour, without having *any particular P.I. at hand* with which to explain it. I think that we have to understand him as talking about the conclusion of a given P.I. which is offered as a hypothesis for the explanation of an action. If there is no particular inference to examine, then we do not have any specific premises and conclusion between which to establish the logical connection.

certain P.I., we need to construct different practical syllogisms which share the same conclusion with our provisional inference, and verify the premises of any one of them.

Consider for example the problem of verifying the proposition 'A opens the door', for which a provisional explanation in terms of A's intention to let X out of the room is offered. We can establish that A's behaviour is an intentional action on his part, only by establishing that the premises of some P.I. with conclusion 'A opens the door', are true. We may for example, construct a hypothetical explanation, which ascribes to A the intention to let Y into the room, and the belief that he could achieve this by opening the door. Then, the verification of the conclusion of our P.I., will depend on the possibility of verifying either one of the conjectured ascriptions, i.e. either that A intends to let X out, or that he intends to let Y in (or indeed, any other premises that conclude in 'A does a'). Hence, the 'burden of verification is shifted from the verification of the conclusion to that of the premises of a practical inference.'[77]

It may be objected here that normally we do not need to verify the intention with which one acts, in order to establish that his behaviour constitutes intentional action, rather than 'mere movements'. For example, we usually know that an agent (intentionally) opens a door, just by seeing his movements; we do not need to establish what is the intention 'behind' his behaviour. This intuition is apparently shared by von Wright himself, since elsewhere he remarks: 'In the normal cases, we say off-hand of the way we see people behave that they perform such and such actions- raise their arms, walk or run, open key-locks or hand things over to one another'.[78] However, these cases can still be covered by von Wright's argument, if he claims that our warrant for 'saying off-hand' that the behaviour we observe is intentional, is that we already have some idea of the intentions that may possibly lie 'behind' it. Hence, we may be able to tell from direct observation that a piece of behaviour is intentional, but this is only because we implicitly assume that there is a further aim and a belief

[77] E&U, p.109.

[78] von Wright [1972], p.32

involved.[79] The ultimate verification of our ascription, would then be again dependent on the verification of the premises of one of the practical syllogisms we are implicitly assuming to hold in this case.

Thus, in the first part of the argument, von Wright tries to show that the verification of *some* P.I. is a necessary condition for the verification of any ascription of intentionality to behaviour. But, does he want to argue that it is also a sufficient condition? This is the question that the second part of the argument deals with.

With it, von Wright tries to establish that the verification of the premises of the P.I., depends on the verification of the conclusion. One possible way that we might attempt to corroborate a hypothesis concerning the agent's motivational structure (i.e. his intention and means/end belief), is negatively : If the agent really intends to bring about p, then he will not, until he fulfils his intention (at the time specified by the premises), embark upon an action which he thinks is incompatible with p. If, for example, A intends to go to the theatre tomorrow evening, then he should not book tickets for a film that is on show at the same time as the theatre performance. If he did that, then we would be justified to doubt whether he really had the intention to go to the theatre, or whether he held the right beliefs about e.g. the time of the performances. This way of falsifying the premises, relies on our ability to construct a practical syllogism that will show that A's booking of the cinema tickets was intentional and, if so, under what description and with what beliefs. It is not possible to overturn the original ascription of intention just by looking at the *results* of A's actions. We have to see the aims and the beliefs that are present, and to do that we have to construct practical syllogisms that explain his actions as intentional. Therefore, this indirect, 'negative' way to verify or falsify a putative ascription of intention still depends on the verification of other intentions and intentional actions.

Other possible methods we might use to establish the truth of the premises of some practical syllogism, are by appealing to the agent's character and to

---

[79] That this is the position that von Wright himself would take, is suggested by the following phrase, quoted from the same passage as in fn.61: 'We are further acquainted with innumerable ends for the sake of which these actions may be performed'.

regularities in his behaviour, considering his educational and cultural background, or taking into account his own self-avowals. Von Wright tries to emphasise that all these indirect ways that help us establish our hypothesis, necessarily involve appeal to the agent's intentions and actions, and hence are themselves based on the interpretative pattern laid out by the P.I.. There is no more direct way of establishing the truth of a given set of premises of a P.I., than by looking at the agent's behaviour and his attitudes through other practical syllogisms.

Consider for example the method of appealing to regularities in the agent's behaviour. The knowledge of these regularities is, according to von Wright, also based on our intentionalistic understanding of the agent's actions on past occasions: Having interpreted his behaviour through a practical syllogism on previous cases, we extrapolate to the present case and make our ascriptions accordingly. Regularities do not take us beyond the agent's behaviour as it is understood through the P.I. schema. Therefore, von Wright claims that it is not possible to appeal to regularities, analogies, character traits, etc. as direct evidence for the existence of a certain intention, because they are also based on intentionalistic, and hence indirect, understanding of the agent's behaviour.

Even the agent's self-avowals may not be considered as direct ways of establishing some attitude-ascription, for they, as well, have to be interpreted by using the P.I. schema. In order to use the agent's declarations (which themselves constitute *verbal behaviour*) as a basis for grounding our hypothesis, we need to understand them as intentional action and explain it in terms of his intentions and beliefs. He may, for example, be lying, or he may be misunderstanding his situation or the meaning of our questions. Again, this kind of verificational procedure provides no more direct evidence for the truth or falsity of our hypothetical premises, than any other appeal to his intentional behaviour.

The point that, as I understand him, von Wright tries to make in the second part of his argument, is that we cannot have direct access to the agent's intentions and beliefs, by somehow getting direct access to his inner states, or by considering some other piece of objective evidence such as 'non-

intentionalistically' understood behaviour. Of course, this is not a claim about our actual scientific or epistemic capacity, but a claim about the logical structure of the concepts of intention and intentional action. The only way we can find evidence for the truth of a certain intention/belief ascription, is by investigating the agent's behaviour as this is mediated by further actions, intentions, beliefs and so on. The intentionality of behaviour is not located at some inner state of the agent, that comes into being at a certain time before the action and somehow (causally or otherwise) brings about the intended result. It has to be established by understanding what is the aim in the agent's behaviour, and this cannot be done by looking for special kinds of 'inner states' behind the behaviour, but by using the P.I. schema to interpret it, whether it is verbal or otherwise.

III.

The problem that von Wright started with, was to establish that a certain ascription of an intention and belief to the agent was true, but in all of the verificational procedures that he considered, there was no direct evidence to be found. All of these indirect methods can be helpful, but they are still fallible and provisional, taking us no further than the agent's intentions and actions. The question is then, whether he should take these indirect methods of verification as acceptable for the establishment of the premises of the P.I., or whether he should say that the only conclusive evidence for them, is that the conclusion of the P.I. be true, i.e. that the action described there is performed. Is the ultimate evidence for the truth of a certain volitional and cognitive attitude ascription to the agent, that he acts accordingly? It seems that, unless such a position is taken, the verifiability argument will be inconclusive as an argument for the logical dependence between the premises and the conclusion of the P.I. schema. For, if von Wright accepts as conclusive evidence the agent's self-avowals, or his past actions and intentions (which are plausible candidates), then it seems that the propositions serving as premises of the practical syllogism under consideration will be verifiable independently of the proposition serving as conclusion, and hence that the argument cannot go through.

As Martin[80] and Tuomela[81] have correctly pointed out, to take such a position amounts to assuming that a logical relation exists. In that case, the argument becomes circular. If von Wright claims that the only factor that would ultimately decide the falsehood of the premises is the non-instantiation of the conclusion, then he is begging the question, since he is simply presupposing that a logical entailment exists, which is what he set out to prove in the first place. Von Wright's argument, if interpreted this way, cannot be considered as proving the logical connection between premises and conclusion from the impossibility of their independent verification, because in order to establish the latter claim, it simply presupposes the former.

IV.

Therefore, von Wright would be caught in a circle if he took the instantiation of the action, as described in the conclusion of the P.I., to be the only conclusive evidence for the truth of the premises and so, the verifiability argument would have no force against the Causal Theories. Nevertheless, we can see his discussion of the verifiability argument as an attempt to *demonstrate* the conceptual dependence between the premises and the conclusion of the P.I., rather than as an argument to exclude the causal connection by establishing that a logical entailment holds. This interpretation is strongly suggested, I believe, by von Wright's discussion of the 'tyrant example', which concludes the exposition of the verifiability argument. R. Martin[82] argues that this example is used as a means to apply 'logical pressure' to the preceding argument, to confine, that is, its domain of application within the limits where the conceptual connection indeed holds.

Under this interpretation, the argument for the mutual dependence of the verification of premises and the verification of conclusions in practical

[80] See Martin [1976], p.336-7

[81] See Tuomela [1977], p.185

syllogisms, would be seen as an attempt to elucidate our concepts of intention and intentional action which, according to von Wright, are such that their explanatory connection is based on a conceptual, rather than a contingent dependence. As he observes at the end of his discussion of the verifiability argument: 'It is a characteristic of these verificational procedures, that they presuppose the *existence* of some factual behaviour, upon which an intentionalist ''interpretation'' is then put'[83]. This phrase confines, as it were, the domain where the conceptual dependence holds. Once we have a piece of 'factual behaviour', we are forced to understand and explain it in terms of some practical syllogism, if we are to see it as action at all. And in our effort to see it as such, no other considerations apart from the ones pertaining to practical syllogisms could help us. For example, there can (logically) be no direct, empirical evidence for the agent's intentions and actions, separate from their intentionalistic understanding, that would help us decide between our putative explanations. So, if the presupposition holds (i.e. we have a piece of factual behaviour on which to apply our explanation), the action featuring in the explanandum is conceptually tied to the premises and it is in virtue of *this* tie that the premises explain the conclusion. The verifiability argument shows that no other considerations are relevant.

But, assume that we lift the presupposition that some 'factual behaviour' is there. 'What does this assumption amount to?', is the question that prompts von Wright to his discussion of the 'tyrant example'. In this philosophical 'thought experiment', we are invited to imagine a case where, all the premises of the P.I. are true by hypothesis, but there is no 'factual behaviour' on the part of the agent upon which to place our interpretation. The specific example given is that of a would-be assassin who is resolved to shoot a tyrant. We are assuming that he has formed the relevant intention and means/end belief, and that all the 'normal conditions' prevail. But, when the time has come to shoot, without having forgotten his intention, he does not shoot, but instead he stands there

---

[82] See Martin op. cit., p.338-9

[83] E&U, p.116

with no relevant behavioural signs whatsoever[84]. After using every possible method of investigation into the case, including medical examination, there are still no signs that show that he changed, or forgot his intention, or that he was prevented by some physical (e.g. paralysis), or psychological (e.g. fear) factor. Assuming then that we have no grounds at all for denying the premises of the P.I., such a case would be a counterexample to the logical entailment thesis, since we have the premises being instantiated as true and the conclusion as false.

The question raised by von Wright is whether we should deny, after all, the truth of the premises, thus turning the conclusion into the ultimate criterion for their truth or falsehood. An affirmative answer, von Wright says, would mean that we have turned 'the validity of the practical syllogism into a standard for interpreting the situation'[85]. It would mean, in other words, that the validity (i.e. the logical conclusiveness) of the P.I. was *assumed*, in order to show that premises and conclusion are not independently verifiable.

V.

Von Wright himself does not opt for the affirmative answer to the question raised above. He says that, 'there is no logical compulsion here.....if this sort of case can be imagined, it shows that the conclusion of a practical inference does not follow with logical necessity from the premises. To insist that it does would be dogmatism'[86]. Hence, his conclusion appears to be that the premises of the P.I. do not logically necessitate the occurrence of an action that matches them. But he goes on to distinguish between the 'prospective' and the 'retrospective' uses of the schema. In its prospective use, the schema is applied for the purpose

---

[84] It is important that he does not appear to be *forbearing to act*, either, for if he were he would be performing a relevant to the situation action (since forbearances are acts for von Wright) on which we could place a new interpretation. That this is so, we can imagine that it is established by questioning him, or physiologically examining him.

[85] E&U, p.117

[86] ibid., p.117

of prediction of action and, as the 'tyrant example' shows, in this case we cannot say that the premises entail the conclusion. But in the retrospective use, where the schema is applied for the purpose of explaining a piece of 'factual behaviour' which is already there, von Wright sees a logical entailment between the premises and the conclusion. This necessity, is necessity 'conceived *ex post actu*'[87] , he writes.

His contention that, in its 'ex post actu' use, the P.I. is logically conclusive, has been, for good reasons, much criticised. To say that in the retrospective use of the schema, if the premises were true, then 'the action was logically bound to happen', cannot mean, as Malcolm[88] correctly points out, that the fact that the agent acted is included as a premise in the P.I.. For if it meant that, then the conclusion would indeed follow from the premises, but it would follow trivially. The fact that the action has taken place, cannot alter the logic of the inference. Since von Wright admits that in the prospective case, the truth of the premises does not entail that the behaviour will occur, he has to concede also that there is no logical entailment involved, whether we stand 'ex ante' or 'ex post actu'.[89]

Von Wright's idea of necessity conceived 'ex post actu' is therefore untenable, if understood as an attempt to establish that a logical entailment, which excludes the possibility of a causal connection, holds between intention/belief and action in the retrospective use of the P.I. schema. However, in the writings that followed 'Explanation and Understanding', and 'On so called Practical Inference' (in which he still defended the idea of 'ex post actu' necessity), von Wright apparently accepted the criticisms raised by Malcolm and others and tried to modify his position. So, for example, in his reply to Malcolm, he wrote that 'the idea to the effect that the practical inference is

---

[87] ibid. P.117

[88] See Malcolm op.cit., p.358

[89] Rex Martin makes the same point in (op. cit., p.341): 'Either we have a logical entailment between the two parts of the inference schema, or we don't:...And if [the premises]...do not logically entail [the conclusion]...then adding the [ex post actu]...stipulation, makes no difference.' (My brackets).

conclusive only ex post actu.....is obscure.'[90] Still, as is the case with all non-causalists, von Wright maintained his intuition that the L.C.A. is substantially right. Nevertheless, I do not think that there is any clear, modified statement of the argument that might prove his case, in his later writings. There are, however, some illuminating remarks made by von Wright there, that might help us establish some sort of coherent version of the Logical Connection thesis that can be used against the Causal Theory.

As I have argued above, von Wright's discussion of the Verifiability Argument, is best seen as an elucidation of the considerations that we take into account when we are trying to establish the truth of an explanation (or prediction) of an intentional action, within the conceptual framework set out by the P.I. schema. The 'tyrant example' shows what the limits of this framework are. The fact that we can imagine a case like that and accept it as a logical possibility, means that the premises of the P.I. do not logically necessitate the occurrence of the action described in the conclusion. And this is true, irrespectively of whether the action has taken place or not. But, within our conceptual framework of action- explanations, accepting the truth of the premises, 'forces' us to accept the conclusion as well. If we do not, then the framework collapses and we can no longer apply its concepts to agents who do not act according to their (verified) intentions and beliefs.

Consider the example of the would-be assassin. The fact that he fails to act, is completely unintelligible to anyone who wishes to employ the notions of intention, belief and action, as they are accommodated within the conceptual framework set out by the P.I. schema. As Malcolm points out, 'if his inexplicable non-performances were other than extremely rare...[then]...we could no longer apply the concept of intention to that person'.[91] On the other hand, when agents act according to their intentions and beliefs, then their actions can be completely explained by means of the practical syllogism. No other kind of consideration (as von Wright has argued in his discussion of the

[90] See von Wright [1984], reply to Malcolm, p.815

[91] Malcolm op.cit., p.364 (My brackets)

verification procedures) would help us understand any better their actions, since whatever further attitudes or behaviour we ascribe to the agent, we do so only through the mediation of practical syllogisms.

The claim then, that can (or should) be defended by von Wright, is that there is a *conceptual* link that holds between the premises and the conclusion of the P.I. and this link can adequately provide for complete explanations of actions in terms of reasons, as long as we keep within the confines of our conceptual framework. Malcolm calls this conceptual relation 'semi-entailment'. The term is in need of clarification, and here von Wright's remarks might help us: In his reply to Malcolm's criticism, he characterises this bond of 'semi-entailment' between premises and conclusion of the practical syllogism, as 'partly a condition of the intelligibility of the action, and partly a condition of the applicability of the concept of intention'.[92]

It is a condition of the intelligibility of the action because, unless the action featured in the conclusion follows from the premises, we cannot understand it as action at all. Having accepted the P.I.-schema as the basis for the explanation of action, it would be impossible to understand how it could be the case that, for an *agent*, the premises were instantiated as true, but the action described in the conclusion did not follow. This possibility cannot be ruled out on logical grounds, but it would have to be considered as a case falling outside our conceptual framework, and thus as completely unintelligible, given the concept of intention that we actually have. On the other hand, the bond of semi-entailment is a condition for the applicability of our concept of intention since, as Malcolm writes, 'we do not ascribe to a person the intention to do a certain thing unless we expect him to do this'.[93] We would not even be able to apply the practical syllogism to agents, if their behaviour (or non-behaviour as in the case of the would-be assassin) could not be explained by it. But we *do* apply this scheme (or so von Wright argues), and we often base our predictions about human behaviour on it, not expecting to come across cases like that of the

[92] In von Wright [1984], p.817

[93] In Malcolm [1984], p.364

'tyrant example'. Moreover, if we do come across one of these 'non-explicable non-performances', we resort to the solution of denying that all the 'normal conditions' prevail, rather than accept that it is genuinely inexplicable.

So, in the light of the above qualifications, we can see, I think, von Wright's discussion of the mutual dependence of the verification of our ascriptions of intention and belief and action, as the spelling out of the implications that follow from our commitment to the P.I.-schema, as the basic schema for the explanation, understanding and prediction of intentional behaviour. These implications stem from the conceptual connection that ties intention to action, and mean that there can be no *direct*, independent from the practical syllogism, verification of our attitude ascriptions. To find the intentionality in the behaviour, we always have to look through the 'glasses' of practical syllogisms.

The Relevance of the Verifiability Argument

I.

But can von Wright's arguments, if construed as suggested above, have any force against the possibility of a Causal Theory of Action? It is a common and well-defended thesis of causalists that no argument supporting the conceptual dependence between reasons and actions, can rule out the possibility of the events instantiating the former being causes of the events instantiating the latter. This view is summarised in the following quotation by J.L. Mackie: 'No problems about independent verification can undermine this possibility of independent existence [of intentions and actions]'.[94]

W.D. Gean[95], has advanced this claim further by arguing that not even an entailment relation between reasons and actions, can exclude the possibility of a Humean causal connection holding between the corresponding events (seen as particulars). Gean accepts that 'if two events have a logical connection, then

[94] Mackie [1974], p.292 (My brackets).

[95] Gean [1975]

they do not also have a causal connection'[96]. Hence, if we can establish the antecedent in the case of intentions and actions, then we will have shown that the former cannot be (part of) the Humean cause of the latter. But, Gean argues that there is no version of the L.C.A., among the ones proposed, that establishes such a connection. At most, what they have shown, is that there is a logical relation between *propositions* referring to the events involved, not between the events themselves. Since the causal relation holds between events and not propositions, the L.C.A. cannot exclude it from the analysis of action.

Consider for example Stoutland's strong type of logical connection, which is defined as follows:

'Two events, A and B, bear the strong type of logical connection to each other, if it is a priori true that when A occurs, B occurs. A and B, for example, bear the strong type of logical connection if the proposition, "If A, then B," is logically true.'[97]

This definition seems ambiguous in a certain respect, since Stoutland presents it as concerning two events, rather than propositions. But, from the way he constructs the definiens, it is clear, I think, that the only interpretations that we can place on this kind of necessity, are either the 'entailment', or the 'de dicto' interpretations. According to the entailment interpretation suggested by Gean, the statement that A and B bear the strong type of logical connection to each other, should be read as follows:

(1): 'Event A occurs' entails 'Event B occurs'.

Gean argues that the necessity involved, if interpreted as above, is compatible with event A being a causal condition of event B. This causal statement should be read as follows:

(2): 'Event A is a causal condition of event B'

---

[96] ibid., p.349

[97] Stoutland [1970], p.119

Statement (1) does not entail the falsehood of (2), since (1) only tells us that a certain *proposition* entails another, whereas (2) tells us that a certain *event* is the cause of another *event*.

The second interpretation that could be placed upon Stoutland's strong type of logical connection, is the 'de dicto' interpretation. According to it, the logical relation involved should be understood as follows:

(3): 'If A occurs, then B occurs' is necessarily true.

Here, the necessity asserted is that of the *truth of a proposition*, not of any relation between the particular events A and B. So, according to Gean, (3) is, like (1), compatible with the truth of (2), and hence 'de dicto' necessity does not rule out a causal dependence.

Gean argues that, the fact that the anti-causalist proponents of the L.C.A. talk about logical connections holding between events only under certain descriptions, amounts to an implicit acceptance, on their part, that the connections in question concern propositions, not events seen as particulars. If those connections held between the events, then we should be able to substitute in them, *salva veritate*, any true descriptions of the intentions, beliefs and actions featured, as is the case for the causal connection. This is admittedly not possible, since it is widely accepted that intentions and beliefs rationalise (and hence are logically connected to) the actions, only under certain descriptions. So, Gean concludes, either a 'de dicto' or an 'entailment' interpretation must be accepted for the necessity argued for by anti-causalists; but such a necessity does not exclude a causal connection. According to Gean, only a statement of 'de re' necessity would suffice to rule out the causal connection. Such a 'de re' interpretation, could possibly be expressed as follows:

(4): There is an event A and an event B, and these events are such that, necessarily, if A occurs, then B occurs.

A 'de re' statement of logical necessity such as the above, would make sure that the logical connection held between the events themselves and not between

propositions referring to them. 'De re' necessity, is therefore incompatible with (2), if the Humean principle of the contingency of causal connections is accepted. But, as Gean points out, even if we disregard the fact that the intelligibility of such statements is a matter of dispute, the anti-causalists have offered no support for the thesis that the connection between the intention and action, is 'de re' necessary. So, he concludes, the L.C.A. as it is expressed by its proponents, does not have any force against the Causal Theory of Action.

II.

Gean's objections against the L.C.A., would appear to be especially telling against von Wright's anti-causal arguments. After all, von Wright has tried to establish a logical connection between the premises and the conclusion of the practical syllogism, between propositions, that is, and not events. So, even if he could establish a strict entailment between premises and conclusion, Gean's arguments would appear to overturn his anti-causal conclusion. Of course, von Wright cannot possibly claim that there is such a thing as a 'de re' necessity holding between intention and action, and being such that the occurrence of the intention is an entity which in itself, logically necessitates the occurrence of the result of the action. Such an entity does not exist since, as Mackie correctly observes: 'Hume's arguments exclude this from reality, not merely from our concept of efficient causation.'[98]

A first point that can be made in defence of von Wright's theory and against Gean's arguments, is that they presuppose a theory of events which is not by any means indisputable and which, in any case, is not shared by von Wright. But, setting this objection aside, we could say that Gean's arguments cannot even be applied to von Wright's Logical Connection thesis, because for von Wright, intentions, and reasons in general, are *not* events at all. Hence, he is not committed to showing that the logical connection in question is a 'de re' necessary connection which holds between events. Therefore, Gean's

[98] In Mackie op.cit., p.295

arguments, if used against von Wright, are question begging, since they rest on the assumption that for every reason that rationalises an action, there corresponds an event. Von Wright of course denies this. The intention of an action cannot be located at some mental state, process or event. It is rather the *direction*, or *meaning* given to behaviour by the agent. In a much quoted extract from 'Explanation and Understanding', he writes:

'One could say - but this too might be misleading - that the behaviour's intentionality is its place in a story about the agent. Behaviour gets its intentional character from being seen by the agent himself or by an outside observer in a wider perspective, from being set in a context of aims and conditions'.[99]

Hence, Gean's 'de re'/'de dicto' distinction cannot be applied to von Wright's arguments, in order to undermine his conclusions. But, by the same token, one could say that von Wright's thesis that intentions cannot be construed as separate occurrences from the behaviour they lead to, is question-begging in its own right against the Causal Theory of Action and therefore, his version of the L.C.A. cannot prove the Causal Theory wrong. Thus, Gean's argument is successful in showing that not even the establishment of an entailment relation can exclude reasons (seen as particular occurrences) from being among the causes of actions, unless the nature of reasons is prejudged in favour of von Wright's 'conceptualistic' theory. I think that this is exactly right , and this is why I contend that von Wright's attempt to establish the L.C.A., if seen as an argument against the *possibility* of a Causal Theory of Action, fails. But, I believe that his discussion of the verifiability argument, is still relevant to the question of whether reasons *are* causes, as opposed to the question of whether reasons *can* be causes. Its relevance becomes clear, if one sees it as a defence of the thesis that a non-causal theory such as his, captures best what is important and genuinely explanatory in our conceptual framework for explaining, understanding and predicting behaviour[100]. This way of looking at the problem

---

[99] E&U, p.115

[100] Of course, one could take the eliminativist position, and claim that there is very little worth retaining in our ordinary psychological, (including action) explanations; we should therefore place our confidence in the evolving science of neurophysiology and its completely different

accords with C. Taylor's position[101], who argues that, we cannot decide on the correct theory on purely 'a priori' grounds: The correct theory can only be decided on its capacity to accommodate the concepts that are relevant to the explanation and prediction of action.

III.

Consider first the relevance of von Wright's arguments to the Causal-Nomological Theory of Action. According to it, the explanation and prediction of action in terms of the agent's intentional attitudes, implicitly assumes causal laws that link the latter with the former. Hempel's insistence that a universal law, or law-sketch, should be included among the explanantia of any adequate explanation, is based on the contention that without any general statement, the explanation would not provide us with adequate grounds for expecting the event described in the explanandum to occur: If there is no universal hypothesis among the premises, then the D-N argument *cannot* carry through.

The practical syllogism, unlike Hempelean D-N arguments, contains as premises *only* singular propositions, which nevertheless suffice to make our expectation that the action will occur, a reasonable one. As von Wright's discussion on the problem of verification has revealed, our concept of intention is such that, only from understanding the meaning of 'A intends to bring about p' and 'A believes that doing q is necessary for bringing about p', we can expect that 'A does q' will be instantiated as true. Our expectation is based on the conceptual dependence between intention and belief on the one hand and action on the other, not on the belief that some causal law linking the two exists. It is true that having established that the agent has reasons to act in a certain way, and that the conditions allowing him to act prevail, we can often predict with

---

taxonomy, replacing little by little today's 'folk psychology'. This is surely a consistent position (taken by e.g. P. Churchland), but it is not one that is espoused by any of the defenders of the Causal Theories of Action which I am discussing here. These theories surely want to retain explanations in terms of reasons, as we ordinarily use them.

confidence the agent's behaviour. In that respect, it may be thought that our predictions resemble the ones made by natural scientists on the basis of causal laws. But as von Wright points out in one of his later works:

'The only hypothetical element involved in a prediction of this kind is that a certain volitional and epistemic attitude of an agent will not change in the timespan allowed between the formation of the intention and the execution of the action. One could therefore also say that the prediction is that the agent will not change his mind, and not that he will act. This is so because of the conceptual, as distinct from the nomic, tie connecting the volitional-cognitive complex with the action' [102].

The generalisations on which we sometimes base our predictions are therefore, statistical generalisations with no causal implications. The only 'risk' of empirical falsification that anyone asserting them takes, is that the agent will not retain his intention and belief until the moment of action. Hence, to construe the schema of practical inference as a law-sketch[103], is to disregard a tenet of Hempel's own theory of explanation: That the criterion of soundness of any explanation, is 'exclusively whether it rests on empirically well confirmed assumptions concerning initial conditions and general laws.'[104] As the verifiability argument shows, there can be no decisive, direct empirical confirmation of the premises or the conclusion of a P.I.; there is always some escape clause suggested by a different interpretation of the agent's attitudes or behaviour, in view of the rest of our ascriptions, which are effected through the use of practical syllogisms. Hence, the power of these generalisations to group together types of intentions and beliefs with the appropriate behaviour, is based on relations of meaning and is always affected by considerations of further intentional attitudes and behaviour, as well as the context within which action takes place. Their conditions of application, in all interesting cases, cannot be fully specified in advance by using direct, (independently of the mediation of

---

[101] See his [1964] book, Ch. 1.

[102] Von Wright [1980], p.56.

[103] As for example Churchland [1970] does.

[104] See Hempel [1942], p. 240.

contextual and intentional considerations), methods of verification. This is why they are statistical generalisations and not sketches of causal laws.

Von Wright's position, as it emerges from his discussion of the verifiability argument, is not so far from Davidson's with respect to his scepticism regarding laws expressible in intentional terms[105]. Davidson's argument against the possibility of establishing strict causal laws in the science of Psychology, rests on his contention that explanations in terms of intentional attitudes are inescapably holistic, and committed to the 'constitutive ideal of rationality'. Von Wright's arguments about the impossibility of independent verification of premises and conclusion of the P.I. schema make use of the same central idea: Any ascription of intention and intentional action *must* be mediated by ascriptions of further intentional attitudes and behaviour. This procedure, is for von Wright inextricably linked with the use of the practical syllogism, whereas for Davidson, it must 'pay allegiance' to the ideal of rationality that we must apply to persons, if we are to understand them as rational agents. Moreover, von Wright's characterisation of the conceptual bond relating intentions and actions as 'partly a condition of the intelligibility of the action, and partly a condition of the applicability of the concept of intention', resonates Davidson's contention that the 'limit...placed on the social sciences is set not by nature, but by us when we decide to view men as rational agents with goals and purposes.'[106] Thus, for both philosophers, the procedure of assigning intentional attitudes must necessarily take into account the holistic relations of meaning that hold between them. A causal law that would be informative and explanatory, would then have to 'allow us to determine in advance whether or not the conditions of

[105] I hope that the preceding discussions of both philosophers' theories, will have made clear the enormous differences that still separate them in almost every other issue in the Philosophy of Action.

[106] 'Psychology as Philosophy', EAE, p.239.

application are satisfied'[107].But because of the objections mentioned above and which both philosophers have raised, there can be no such law[108].

However, there are philosophers[109] that have argued that holism, as well as the allegiance to an 'ideal' principle can also be encountered in physical theories that can nevertheless, still formulate causal laws. They support their contention, by analogies, such as Newton's Mechanics, or the principle of energy conservation. It seems to me though, that the analogy is not warranted. In the Physical Sciences, strict causal laws can be tested in interesting situations, because we can determine in advance whether or not they hold in our experimental set-up. And although we cannot build an ideal closed system, we can often approximate it very closely and experiment in it. In the study of human behaviour, we cannot even approximate any interesting closed system in practice. How can we ever be sure what are the determinants of an agent's intentions and beliefs? If we are to test any interesting actions that resemble real life ones, there is no way that we can know what and how determines the agent's volitional and cognitive attitudes. I think that the arguments by Davidson and von Wright which have already been presented, have made the possibility of establishing beforehand the conditions of application of such laws, look remote.

Hence, if the conditions in which the law holds are not specifiable in advance, then to insist, 'ex post actu', that some causal law that covered the case exists, seems to be an idle move, and can have no force as a defence of the Causal-Nomological Theory. As von Wright observes, 'not the law itself, but the existence of a law is now a tautologous after-construction'.[110]

---

[107] Ibid., p.233.

[108] For a similar position on the affinity of von Wright's and Davidson's views on this matter, see Stoutland [1982], pp.60-62.

[109] See, e.g. Churchland [1970] and Crane and Mellor [1990].

[110] Von Wright [1981].

IV.

The arguments advanced above against the Causal- Nomological Theory do not bear directly on Davidson's Causal Theory, since he does not rest his causal condition on the nomic connection between intentional attitudes and behaviour. Does von Wright's verifiability argument have any force against Davidson's theory? It is difficult to argue against Davidson's theory from the 'intentionalist' standpoint, since Davidson appears to accept all that is considered essential to action-explanations by von Wright; he agrees that the relation between reason and action is not simply empirical and that, as we have seen, there can be no more than statistical generalisations linking the two. So, most of his arguments are geared towards establishing the *possibility* of reasons being causes, rather than the *necessity* of this being so; his only argument in support of the view that the idea of cause must be included in the framework of action explanations, is that only causal explanation can give an adequate account of how the event described in the explanandum is brought about. As I have already tried to show, the thesis in favour of the possibility of a causal connection between reasons and causes, cannot be overturned by any version of the Logical Connection Argument, developed in the context of von Wright's theory, or indeed in that of any 'intentionalistic' theory. Even if a strict entailment between premises and conclusion of the P.I. was established, it would not follow that reasons cannot be non-nomic causes. This is something that Davidson's as well as Gean's arguments have established.

As far as Davidson's thesis that reasons *must* be causes is concerned, I think that von Wright's verifiability argument can be used to undermine it, in the same way that Davidson's own arguments in favour of the anomalism of the mental can be used to undermine the causal relevance of intentional attitudes in Davidsonian action-explanations. As I have argued in Part I, the causal condition in Davidson's model of action-explanations does not help us to understand any better how an *intentional action* (rather than an event *which happens to be* an intentional action), was brought about. The causal condition is simply irrelevant to the explanatory adequacy of reasons, since no account can be given for the causal role played by the *content* of intentional attitudes in

bringing about the action. Von Wright's position is that the conceptual connection between reasons and actions suffices to account for the production of behaviour and that no restatement, in causal terms, of teleological action explanations, is needed. Thinking of reasons (or the onset of reasons) as separate occurrences in the agent's brain that cause his bodily behaviour (which can thus be described as action), without being able to connect (directly or indirectly) the two events *as described* , via a causal law, does not help us understand, explain or predict the action any better. These three purposes are best accomplished by the P.I.-schema without the need of an additional causal clause. The only purpose accomplished by this addition (as effected in Davidson's theory), is to help satisfy the prejudice of anyone who thinks that the relation of event causation must somehow underlie all phenomena that are subject to explanation and prediction. An implication of this line of thinking, is that no explanation can be adequate, if it does not mention the event-cause of the event described in the explanandum. This, I think is the causalist conviction that lies behind Davidson's main argument in favour of the view that reasons must be causes. In the next chapter, I will examine the argument, and consider a possible response by von Wright.

CHAPTER 6 : TWO CRITICISMS OF THE THEORY

The Problem of Justification vs Explanation

Davidson's main reason for holding that action *must* be analysed in terms of causes, is that only the causal condition can account for the explanatory, as distinct from justifying, force of reasons; in non-causal theories, the efficacy of reasons in producing the behaviour involved in the act is simply not there. Such theories, construe the relationship between the agent's volitional and cognitive attitudes and his behaviour as a conceptual one, and hence can only provide rationalisations (which are for Davidson adequate only for justification), but not explanations of actions. This causalist charge is put succinctly by Mackie, when he says that in von Wright's account 'we have nothing but a set of interlocking concepts, with no shadow of an account or explanation of how A *came to* do B'.[111] To explain an action, we need to say, not only in what way it appealed to the agent, but also what brought it about, how it was produced.

Davidson brings out the significance of the causal factor, by producing the (consistent) hypothesis of agents holding more than one reason for the same action, but performing it because of *only* one of them. Without causation, he claims, we cannot distinguish between an agent's acting *because of* a reason and his acting *and* merely having a reason. Consider the following example that illustrates the problem: Suppose that A wants both to let some air into the room, and to let a fly out, and he thinks that the means appropriate for each end, are to open the window. Imagine however, that he opens the window *only* because of his desire to air the room and *not* because of his desire to let the fly out. This does not mean that he is free from his desire to let the fly out, only that he does not *act on* it. When he acts, he desires both the cooling of the room and the fly to be out and therefore his action can be rationalised by both desires. But he

[111] Mackie [1974], p.292.

only acts because of the first one, the second desire merely 'being there', and so we can only explain his action by reference to the first one, not the second. Now, this kind of case is conceivable, if not actual, and is produced by Davidson in order to show that, without the concept of the cause, we cannot distinguish explanation from justification.

I have argued in Part I, that the causal condition, as construed by Davidson, does not give us any help in understanding the efficacy of reasons. Here, I want to examine what kind of response can be given by von Wright, independently of the arguments raised against Davidson.

Von Wright's construal of this kind of cases is completely different from the causal theorist's and is capable, I believe, of meeting Davidson's charge of explanatory inadequacy. In von Wright's theory, the relation between reason and action is construed in a way that does not allow for Davidson's criticism to even be raised. Consider the example given above, seen after the action has been performed, when we seek an explanation of A's opening of the window. Saying that A's behaviour was performed in accordance with a particular intention (e.g. to let in some air), means that this intention was the reason *for* which A acted as he did. If A also possessed an intention to let a fly out which remained inert in that case[112], then this was not the reason for which he acted, because it was not the 'ousting' of the fly which was intended by A's opening of the window. In other words, von Wright's model for the explanation of action, requires that the intention which is given in the explanation of the action, is the

[112]My reference to an 'inert' intention in the context of von Wright's theory, requires qualification, since Von Wright differentiates between intentions for immediate action and intentions for action in the future. He denies that an agent can have an intention to act *now* which remains inert in the above sense: It is a contradiction to say that A intended to p and thought q'ing *right then* necessary for p, but he q'ed for a different reason. However, he allows for a case where A intended to p and thought q necessary for p (but not at that particular instant), and the intention to p remained inactive until he q'ed for a different reason. (See von Wright [1980], p. 59) It is this kind of cases that my defence of von Wright's theory concerns, and it seems to be the only kind which could instantiate the causalists' imaginary examples. Whether von Wright is right in holding that the first kind of cases are not logically possible, is a complex issue which I will not examine here. For a similar to von Wright view, see Ginet [1990], p. 145.

intention with which the agent acted. Therefore, when we explain an action within this model by stating the agent's intention and belief, we are giving its determinants, i.e. the reasons for which the agent acted, not some rationalising attitudes that he simply possessed. An explanation based on the P.I. schema, is explanation in terms of *these* reasons; if not, then it is simply a wrong explanation, based on a fallacious understanding of the intention which was involved in the action.

Davidson contends that the idea of acting for a reason is obscure unless we understand it as implying the presence of a causal connection between reason and action[113]; the only plausible genuinely explanatory relation between the two, is the causal one. But as we have seen, if we include the causal condition in the explanation of action in the way that Davidson suggests, we lose the explanatory relation between the content of the reason and the action. In von Wright's theory, the explanatory relation is solely the conceptual connection, while the notion of agency cannot be analysed any further. The fact that there are agents that intentionally bring about events in the world is as basic, for him, as the fact that there are events that causally necessitate the occurrence of other events. I will not go into this dispute here, since I think that persuasive arguments for either view, can be given only from the standpoint of a comprehensive theory of one or the other kind. All I want to argue here is that, if the notion of agency is accepted, then the distinction between acting with a certain intention and acting and merely having this intention, can be sustained in von Wright's theory.

So, what about the objection raised by the causalists, i.e. that von Wright's theory cannot distinguish between the intention with which the agent acts, and the intention that he merely possesses, but remains inert? I believe that this objection is unwarranted. If by distinguishing between intentions (or, for that matter, desires and other intentional attitudes), the causalist means locating an antecedent state, or event that somehow (causally or otherwise) brings about the behaviour in the act, then he is misconstruing von Wright's theory. The theory

[113] See e.g. ARC, in EAE, p.9.

cannot provide conditions for locating intentional attitudes construed as above, since, by its nature, it does not accept the event-causal analysis of action.

But, if the issue is not prejudged by formulating the problem in event-causal terms, then von Wright's theory can indeed allow for conditions of distinguishing between the intention with which one acts, and the intention that stays 'inactive'. These conditions will certainly not distinguish between a certain antecedent event that motivates the behaviour and another one that does not. The conditions for the correct ascription of the motivating intention 'behind' behaviour will involve appeal to what the agent aimed at, or meant, by his behaviour. To determine that, we will have to look at the agent's character, his future and past behaviour, the particular circumstances, and so on. For example, if A opened the window with the intention of airing the room and not with the intention of letting the fly out, we could probably determine that by looking at his subsequent behaviour. (e.g. he may breathe with a sigh of relief when the window is opened while showing no interest in the presence of the fly.) So, these conditions will be based on the complex considerations of context and ascription of intentional attitudes, discussed by von Wright in his verifiability argument, which show whether the agent intended *by his behaviour* to achieve one end or another. These considerations are bound to be fallible, but no more so than any consideration pertaining to ascriptions of intentionality.

The Problem of Congruence

In view of the above considerations, I conclude that the charge made by Davidson against non-causal theories, that they cannot distinguish between explanation and justification of action is, at least for von Wright's theory, ungrounded. However, if the argument is understood as an attempt to show that the theory cannot explain how 'mere behaviour', i.e. the bodily movements involved in the act, come about, then it has some force and must be considered on a different basis. Von Wright's theory says nothing about the question of

how 'mere' behaviour is produced. His P.I. schema, is a schema for the explanation of action, not bodily movements, and thus when a piece of behaviour is explained according to it, it is explained as an action, there being no direct implications about how the muscular activity involved was produced. In Davidson's theory, on the other hand, actions *just are* bodily movements with the right sort of causes. It is therefore easy, when contrasting the two theories, to equivocate the question of how *'mere behaviour'* is produced with the completely different, for von Wright, question of how *intentional behaviour* is produced. This equivocation, which von Wright does not accept, gives an initial plausibility to the causalists' claim that von Wright can not explain how *actions* come about. This claim is, as I have argued, false. Nevertheless, the question of how 'mere behaviour' is produced, remains unanswered in his theory, hence raising for him as well the problem of congruence. In this section, I will examine in exactly what the problem consists for von Wright, and try to show that it is a real one for his theory.

I.

The main thrust behind the Causal Theory of Action is that it promises to bridge the chasm that seems to exist between the category of agency which is central to our understanding of phenomena that involve human interference, and the category of event causation which is supposed to be the basis of our understanding of the physical world, the 'cement of the universe', as it was called by Hume. The reconciliation of human agency with the naturalist perspective is supposedly achieved by the C.T.A. through an analysis of action that identifies the causes of behaviour, with the mental attitudes that rationalise it. Hence, the only dependence relation that is presupposed is the relation of event causation, and the only type of explanation used is causal explanation; intentional action is thus analysed into categories well-recognised by the physical sciences.

As I have argued in Part I, Davidson's theory does not succeed in giving a satisfactory account of how this reconciliation is made possible. The physical language, which contains the predicates featuring in the causal law that grounds the explanatory causal connection, is radically detached from the psychological language which contains the mental predicates that rationalise the action. This detachment does not let us have any explanation of the relevance of the content of our intentional attitudes to the occurrence of a certain type of events, which are called actions, in the real world.

On the other hand, the Causal-Nomological theory of action does not seem to face this kind of problem. Reasons are nomically tied to their effects and therefore, the causation of behaviour is subsumed under a full-blown causal law that connects mental attitudes *qua mental* with behaviour. So, the answer given by this type of theories to the question why did a particular mental attitude cause a particular piece of behaviour, is because a causal law covered this case. The fact that the existence of a causal law connecting reasons and actions, *so described*, is asserted, is sufficient to justify our choice of describing the relata of the singular causal relation in that way; our need to view content as playing an efficient role in the production of action, is satisfied. The further question of how these mental properties can be instantiated in, or reduced to, physical ones, belongs to the domain of the mind/body problem and does not affect the integrity of the causal-nomological explanation of action. However, powerful arguments have been developed against the causal-nomological theory by causalists and non-causalists alike so that the theory is shown to be at odds with the way we use our intentional concepts. The Causal-Nomological theory, for that reason, cannot be considered as providing a plausible account of intentional action, despite the fact that it appears to deal with the problem of congruence in a satisfactory way.

II.

Von Wright's theory can also been criticised[114] on the grounds that it fails to provide an explanation for the congruence between action and the bodily movements that are necessary for the accomplishment of its result. Our demand here, is for an explanation of the fact that my action of q-ing, which is not to be analysed in event-causal terms, is regularly accompanied by an event in the world which is intrinsic to my action, and which is obviously the effect of some neurophysiological cause. In the case of (productive) basic actions, this event is some bodily movement. Now, in causal theories, the cause *equals* the reason, and the effect *equals* the action. In his theory, von Wright has constructed a schema for the explanation and understanding of intentional action, but he has not provided an account in the same terms for the occurrence of these bodily movements (or, in general, of the events) that are necessary for such actions. Hence, the question naturally arises: How can we explain the fact that, whenever I act intentionally, the event which is intrinsic to my action occurs as an effect of some (sufficient) causal condition?

III.

At a first level, Von Wright tries to meet this objection by denying the need for an explanation. The explanandum of a teleological-intentionalist explanation is different from that of a causal explanation. To explain causally a piece of behaviour, is to give sufficient causal conditions for the occurrence of some bodily movements. This will presumably involve a description of a neurophysiological structure involved, as well as the 'triggering' neural events that produce the movements. But, no matter how detailed the explanation is, it will necessarily refer to the movement of the agent's body, not to the agent's

[114] See Stoutland [1982], [1984].

movement of his body. Hence, the explanandum of a causal explanation of behaviour must be a physical event that makes no mention, implicit or explicit, to agency.

On the other hand, the explanandum of a teleological-intentional explanation of behaviour, is an action. According to von Wright, a P.I. cannot be applied to explain behaviour without having first placed an intentional interpretation on it. To identify and describe appropriately an explanandum, one has to first understand it as intentional action: 'In order to be teleologically explicable....behaviour must first be intentionalistically understood[115]. In the process of picking out the descriptum which we need to describe and have explained, it is not our observation of physical movements 'stripped' from intentionality which matters, but our 'intentionalistic' understanding of the behaviour as action. This is evident especially in actions that involve a multitude of complex and disparate movements which can be grouped together and subjected to explanation, only in virtue of their being movements that constitute a such and such intentional action. Hence, to describe in the appropriate way and consider as an explanandum an action such as enrolling on a three year course of Biology, say, we first have to understand the great variety of movements involved as being intentional; no considerations of their physical aspect will help us. Once we have identified and described our action-explanandum, its physical aspect is immaterial for the purposes of the teleological explanation.

It is Von Wright's contention that when we observe an action, we do not observe movements of the body and then place on them an intentionalistic understanding. The two forms of description are rather two different ways of understanding something (behaviour) which is initially neutral. The physical form of description involves grasping what behaviour is like in certain physical respects, whereas the intentional form of description involves grasping the meaning of behaviour. In both cases, understanding in one way or another takes place before explanation can take off: 'Understanding is a prerequisite of every

[115] E&U, p.121.

explanation, whether causal or teleological'[116]. The 'reading off' of actions from behaviour is as direct as the observation of physical movements of the body, and in the case of action-explanations, it is *this* kind of understanding that matters. Von Wright thinks that it is false to believe that our understanding of action involves the placing of a certain interpretation on bodily movements. What we see is a person raising her arm, not a person's arm rising which we *then interpret* as her raising of her arm.

Therefore, it is a mistake, according to von Wright, to think that the behaviour in the act is 'mere behaviour' (i.e. physically described), which is given an intentionalist interpretation. If this is so, then no explanation can be asked for the fact that the behaviour in the act occurs congruently with the action. For the behaviour in the act is not bodily movement but intentional behaviour, and as such, it cannot have a causal explanation at all. Once we have a physical description of behaviour, a causal explanation is possible; likewise a teleological explanation is possible when we start with an intentionalistic description. But since the latter does not contain in any sense the physical description as its basic constituent, the question of how the behaviour involved in the action can also have a causal explanation is illegitimate: These are explanations of two different things.

IV.

But this way of denying the need for an explanation of the congruence between bodily movements and action, will not satisfy us when we inquire about the congruence between actions and their results. The result of an action is an event (a 'change in the world') which, being an intrinsic part of the action, its occurrence is explained through the intentional-teleological explanation: It is a change in the world brought about by the agent for some particular reason. But it cannot be disputed (except perhaps by agent-causalists) that this change is also the end-effect of some event-causal sequence that (presumably) begins in

[116] E&U, p.135.

the agent's brain and, after proceeding through her nervous system and muscles, produces the result of the action. That this event can have two different types of explanation, is something that von Wright has to accept, if his theory is not to defy well-established scientific facts. His task must then be to show that the two types of explanation are compatible, and furthermore, that any demand to the effect that they be somehow connected, is illegitimate.

Von Wright approaches the problem by considering the counterfactual element involved in action.[117] He imagines a case where the result of an action would have materialised independently of the agent, and asks whether in such a case we would deny that the agent had really performed an action. For example, we are invited to imagine her opening intentionally a door by turning the handle and pushing, while at the same time, unknown to the agent, there is a causal mechanism operating and opening the door just at the time that she is pushing. In that case, von Wright says that we should deny that the agent *opened the door* since, the result of this disputed action was not brought about by the agent but by an independent cause operating without her knowledge. So, the counterfactual statement, 'The result would not have materialised, had she not performed the action', is false, forcing us to retract our original ascription of intentionality to the agent's behaviour. Still, we could redescribe her behaviour as an act of 'trying to open the door' and this would impute intentionality to it, even if it would be only 'mutilated' intentionality.

However, in the case of basic actions no such redescription is possible, since basic actions are not performed by doing something else. If I fail to raise my arm (as a basic action) I cannot be said to have tried but failed to raise it: I either can raise my arm, or I can't. So, if we imagine a causal mechanism of the sort envisaged in the previous example, operating independently of the agent whenever she raises her arm, then we cannot redescribe her action as a *trying* to raise her arm. If we are to deny that *she* raised her arm, then we cannot impute on her behaviour a 'mutilated' aspect of intentionality, because if *she* did not raise her arm, then *she did nothing.*

---

[117] This discussion is found in E&U, p.125 ff.

But, according to von Wright, the fact that the counterfactual 'if the agent had not raised her arm, then the arm would not have risen' is false, should not lead us to deny the intentionality of such basic actions. He thinks that the result of a basic action, (e.g. the arm's rising) can have a Humean event-cause that nomically necessitates it, and still be the result of an action, i.e. its intrinsic event that is brought about intentionally. The occurrence of the event-result can thus be explained both by subsuming it under causal laws, and teleologically, as the intrinsic event of an intentional action. This defence of compatibilism, amounts to denying that the kind of objective counterfactual element involved in non-basic actions, is also present in basic actions. So, to say that the result of the basic action would have occurred independently of the agent's intentionally bringing it about, does not imply that the agent did not perform an intentional basic action. However, von Wright still retains the counterfactual element for basic actions, but only in a subjective sense. 'The element of counterfactuality' he writes, 'consists in that the agent *confidently thinks* that certain changes will not occur unless he acts'.[118] So, if the agent believes that certain causes operate independently of him, and he just observes them producing results such as risings of his arm, then he necessarily *does not* act. This is so because to observe a cause operating, implies that you stay passive: 'When I observe, I *let* things happen. When I act, I *make* things happen'.[119]

I think that von Wright's way of arguing for the compatibility of the two types of explanation is not convincing: He does not make clear what is the relevant difference between the example of the non-basic action and that of the basic action, that leads him to deny the intentionality of the first one, while affirming the intentionality of the second[120]. As a consequence, his denial of the objective counterfactual element in basic actions, looks artificial as a solution to his problem.

---

[118] E&U, p.130, fn.39.

[119] ibid., p.130.

[120] Perhaps he means to treat the first case as a case of pre-emption, and the second one as a case of overdetermination; but this is not made clear in his discussion of the examples. Moreover,

Nevertheless, I believe that von Wright's theory indeed allows for the two different types of explanation being compatible in the case of basic actions. To see that, we need to specify a criterion for the incompatibility of different explanations, such as Kim's general principle of explanatory exclusion, which states that, 'two or more *complete* and *independent* explanations of the same event cannot coexist.'[121] In the problem that we are considering, we have two explanations of the same event (the result of the action), which are indeed independent (since von Wright does not allow for any connection between the two), but not both of them are complete. The causal explanation of the rising of A's arm is complete since, we are assuming, there must be antecedent sufficient conditions for its occurrence, in terms of some neurophysiological events and mechanism. However, the explanation of the same event in terms of A's intentional attitudes, is not complete, since von Wright's schema is a schema for the explanation of *actions,* not of the *results of actions*. So, when we explain A's arm-raising by citing her relevant intention and belief, we are explaining the result of her action only indirectly, since our explanandum is an instance of a *relation* that holds between A and the rising of her arm. Therefore, even if we consider the explanation of the action as complete, the explanation of its result need not be complete. We can see that it is not complete in the sense that the causal explanation is, if we notice that the occurrence of *this particular* arm-rising is not necessitated by A's attitudes, as it is necessitated by its sufficient causal conditions: There is no direct reference to any particular arm rising in the intention or the belief cited in the explanation.[122] The teleological explanation

---

this would be a misleading way to characterise the situation, since in the case of the basic action, the result is not *caused* by the agent.

[121] See Kim [1989], p.89. (My emphasis). Kim restricts the use of his principle to causal explanations, but then he goes on to apply it to Malcolm's model of action-explanations (see Malcolm [1968]) which is, according to Malcolm, irreducibly teleological. The source of the confusion is the fact that Malcolm uses the word 'cause' not in the Humean sense, but in its widest sense, meaning 'anything that explains or partly explains the occurrence of some behaviour.' (ibid., p.59) In that sense, von Wright's explanantia, can also be called causes of behaviour, and hence the principle can be applied to explanations in terms of them.

[122] For a way in which (concurrent) intentions, involving direct reference to their object (i.e. the action), can be construed, see Ginet [1990], p.136ff.

is, however, a complete explanation of the action, i.e. of *A's bringing it about that her arm rises.* Its completeness does not, of course, consist in the fact that A's attitudes nomically necessitate the action, but in the conceptual connection that holds between the two. Hence, I conclude, the teleological, indirect explanation of the result of a basic action, is compatible with a complete causal explanation of the same event[123].

V.

The arguments presented above, establish the compatibility of the two types of explanation, but they do not yet solve the problem of congruence. There is an inherent instability in a situation where we have, for the same event, two compatible explanations which are supposed to hold in virtue of two different explanatory relations. In our case, we have a teleological, incomplete explanation of the event in question, which, according to von Wright, holds in virtue of a relation between the agent and the result of her action, the relation of *bringing about.* This relation neither is, nor can be analysed in terms of, the causal relation. For von Wright, it is just part of the way the world is, that there are agents who bring about changes in the world. On the other hand, we also have a causal explanation of the result of the action which, we believe, is grounded on a causal connection between a neural event and the result itself. Explanatory realism requires that this causal connection is objective, i.e. that its instantiation is independent of our explanations in terms of it.[124] For anyone who wants to view the possibility of there being agents that intentionally bring about events in the world, as a real one, there is a question that needs to be

[123] Of course the conclusion holds only with respect to Kim's principle of explanatory exclusion (which gives us only a sufficient condition for rendering two explanations incompatible), but I think that the principle captures what, intuitively, seems to be the source of the problem here: That there appear to be two independent, complete explanations for the same event.

[124]With respect to the 'objectivity' of the causal explanatory relation, I am here using Kim's idea of explanatory realism (see Kim [1987] and [1989]). With respect to the relation of 'bringing about', I do not think that we can call it objective in the same sense; it suffices I think to say that von Wright considers it an irreducible relation.

answered here: How can it be the case that, whenever I perform an action by making it true that a certain change in the world is effected, that very change is also causally necessitated by some antecedent neural event? There is a congruence between my bringing about of certain events, and nature's causal mechanism that produces the same particular events, and this congruence seems surprising. Not to von Wright though. For him, the fact that the two different types of explanation may both be valid on every particular occasion, is a contingent fact; it is just part of the way the world is, and thus defies any explanation:

'That this requirement is met, is a contingency. But it is nothing to be surprised at. For, it is a condition which the world must satisfy, if we are to entertain our present notions of action and agency.' [125]

So, his response with respect to the challenge to account for the congruence between action and event-causation, is to deny the legitimacy of questions demanding that an explanation should be given. Such an argument is presented in his *Causality and Determinism.*[126] There, it is argued that for the establishment of each type of connection, there are sufficient conditions, and these two sets of conditions are independent of each other. Consider the case of an agent getting up from a chair as a basic action. If we want to establish that the agent brought about the event involved (that his body got up from the chair) as a result of a basic action, we have to make enquiries about his intentions, abilities, awareness of the situation, etc. The existence of an independent cause will be established on the basis of a completely different form of enquiry, an enquiry into the physiological state of the agent. But these two types of enquiry are independent of each other (at least when the agent has the physical ability to perform the action) and so, there cannot be any influence of the conclusion of the one on the conclusion of the other. Even if we establish that a sufficient cause for the lifting up of the man from the chair was operating, we could still maintain that the agent really got up from the chair intentionally, provided of course that such a conclusion is suggested by our investigation into his

[125] von Wright [1974], p.132.

intentions, beliefs, abilities, etc. The two forms of enquiry are therefore compatible but also detached, so that no explanation for their congruence can be given.

If we attempted frequently to perform a certain type of action without its result materialising, or if the event constituting its result occurred unintentionally (e.g. unintentional arm-risings), on many occasions, then we would lose confidence in our ability to perform this particular type of action. And if these discrepancies between 'mere behaviour' and action occurred for all types of action, then we would probably give up the concept of intentional action altogether.

'The events which are the results of basic actions thus happen, on the whole, only when we «vest» these events with intentionality, i.e. perform the basic actions. That this should be so is an empirical fact, but a fact which is fundamental to the concept of an action. The conceptual basis of action, one could therefore say, is partly our ignorance (unawareness) of the operation of causes and partly our confidence that certain changes will happen only when we happen to be acting.'[127]

It seems to me that von Wright's response is a poor one, for the same reasons that Davidson's rejection of the need for explanation of the congruence is unsatisfactory. The fact that a necessary constituent of my basic actions, their intrinsic result, occurs as an effect of some event-cause, whenever I intentionally bring it about, is indeed something that must be explained by a theory that purports to provide a complete account of action. The problem is, that this event is a necessary part of the action itself: There cannot be an arm-raising without an arm-rising. So, if the agent's bringing about of this event, remains detached from the type of connection that is justifiably considered as the most appropriate explanatory connection for its type, i.e. the causal connection, then it must be the role of the 'actionist' relation that is brought under suspicion. Our intuition is more inclined to lead us to question the explanatory role of intentionality in the 'production' of physical events, than to

---

[126] ibid., p.132.

[127] E&U, p.130.

accept that it is a mystery how brain events, congruently with agents, bring about the results of their actions. Therefore, if we were to accept von Wright's contention that the role of event-causation is independent from that of the agent in bringing about these events, then we would be drawn towards a non-realist understanding of intentional explanations.

My aim is not to question von Wright's metaphysical assumptions, which include an 'actionist' understanding of causation, but to argue that his theory, like Davidson's, cannot provide explanations of action which are complete, in the sense of answering all genuine questions which concern the relevant, for the purposes of explanation, aspects of action. The question of the occurrence of the result *is* a relevant question, since it concerns a logically intrinsic part of action. His admittance that physical causation may be at work here, puts us in a predicament of wondering at the fact that two types of relation, which are radically detached, are both responsible for the occurrence of one and the same event. As he, himself says, 'wonder of this type can be both the starting point and the end station of philosophical inquiry'.[128] For his theory's sake, I hope it's not the latter.

[128] See his reply to Stoutland, von Wright [1984], p.809.

BIBLIOGRAPHY

Anscombe, G. E. M. [1984], 'Von Wright on Practical Inference', in Schilpp (ed.), *The Philosophy of Georg Henrik von Wright*, The Library of Living Philosophers, pp. 377-404.

Antony, L. [1989]. 'Anomalous Monism and the Problem of Explanatory Force', in The Philosophical Review 98, pp.153-87.

Bishop, J. [1989]. *Natural Agency*. Cambridge University Press, Cambridge.

Chisholm, R. [1966]. 'Freedom and Action', in *Freedom and Determinism*, K. Lehrer (ed.), Random House, New York, pp. 11-44.

Churchland, P. [1970]. 'The Logical Character of Action-Explanations', in The Philosophical Review 79, pp. 214-36.

Crane, T. and Mellor, D. H. [1990]. 'There is No Question of Physicalism', in Mind 99, pp. 185-206.

Cummins, R. [1983]. *The Nature of Psychological Explanation*. The MIT Press, Cambridge.

Davidson, D. [1980]. *Essays on Actions and Events*. Clarendon Press, Oxford.

Davidson, D. [1987]. 'Problems in the Explanation of Action', in Pettit, P. et al. (eds.), *Metaphysics and Morality- Essays in Honour of J.J.C. Smart*. Oxford.

Davidson, D. [1993]. 'Thinking Causes', in Heil, J. and Mele, A. (eds.), *Mental Causation*. Clarendon Press, Oxford, pp.3-17.

Dray, W. [1957]. *Laws and Explanation in History*. Oxford University Press, Oxford.

Dray, W. [1963]. 'The Historical Explanation of Actions Reconsidered', reprinted in Gardiner, P. (ed.), *The Philosophy of History*. Oxford University Press, Oxford (1974), pp. 66-89.

Ducasse, C. J. [1925]. 'Explanation, Mechanism and Teleology', in Journal of Philosophy 22, pp.150-5.

Gean, W. D. [1975]. 'The Logical Connection Argument and De Re Necessity', in American Philosophical Quarterly 12, pp. 349-54.

Ginet, C. [1990]. *On Action*. Cambridge University Press, Cambridge.

Hempel, C. [1942]. 'The Function of General Laws in History', reprinted in Hempel, C., *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. Free Press, New York (1965), pp. 231-43.

Hempel, C. [1963]. 'Reasons and Covering Laws in Historical Explanation', reprinted in Gardiner, P., (ed.), *The Philosophy of History*. Oxford University Press, Oxford (1974), pp. 90-105.

Hempel, C. [1965]. 'Aspects of Scientific Explanation', in Hempel, C., *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. Free Press, New York, pp. 331-496.

Honderich, T. [1982]. 'The Argument for Anomalous Monism', in Analysis 42(1), pp.59-64.

Kim, J. [1987]. 'Explanatory Realism, Causal Realism, and Explanatory Exclusion', in Midwest Studies in Philosophy 12, pp.225-39.

Kim, J. [1989]. 'Mechanism, Purpose, and Explanatory Exclusion', in Tomberlin, J. (ed.), *Philosophical Perspectives 3: Philosophy of Mind and Action Theory*. Ridgeview, California, pp. 77-108.

Mackie, J. L. [1974]. *The Cement of the Universe*. Oxford University Press, Oxford.

Malcolm, N. [1968]. 'The Conceivability of Mechanism', in The Philosophical Review 77, pp.45-72.

Malcolm, N. [1984]. 'Intention and Behaviour', in Shilpp (ed.), *The Philosophy of Georg Henrik von Wright*, The Library of Living Philosophers, pp. 353-76.

Martin, R. [1976]. 'The Problem of the «Tie» in von Wright's Schema of Practical Inference: A Wittgensteinian Solution', in Hintikka, J. (ed.), *Essays on Wittgenstein in Honour of G. H. von Wright*, in Acta Philosophica Fennica, 28, pp. 327-63.

Melden, A. I. [1959]. 'Willing', in The Philosophical Review 69, pp. 475-84.

Nickles, T. [1977] 'Davidson on Explanation', in Philosophical Studies 31, pp.141-5.

Ruben, D.-H. [1990]. 'Singular Explanation and the Social Sciences', in Midwest Studies in Philosophy 15.

Stoecker, R. [1993]. 'Reasons, Actions And their Relationship', in Stoecker (ed.), *Reflecting Davidson. Donald Davidson Responding to an International Forum of Philosophers*. Walter de Gruyter, New York, pp. 265-86.

Stoutland, F. [1970]. 'The Logical Connection Argument', in American Philosophical Quarterly Monograph 4, pp. 117-29.

Stoutland, F. [1976a]. 'The Causal Theory of Action', in Manninen, J. and Tuomela, R. (eds.), *Essays on Explanation and Understanding*. Synthese Library, Reidel, Dordecht and Boston, pp. 271-304.

Stoutland, F. [1976b]. 'The Causation of Behaviour', in Hintikka, J. (ed.), *Essays on Wittgenstein in Honour of G. H. von Wright*, in Acta Philosophica Fennica, 28, pp. 286-325.

Stoutland, F. [1980]. 'Oblique Causation and Reasons for Action', in Synthese 43, pp.351-67.

Stoutland, F. [1982]. 'Philosophy of Action: Davidson, von Wright, and the Debate Over Causation', in *Contemporary Philosophy: A New Survey, Vol. 3*. Martinus Nijhoff, The Hague, pp. 45-72.

Stoutland, F. [1984]. 'Von Wright's Theory of Action', in Schilpp (ed.), *The Philosophy of Georg Henrik von Wright*, The Library of Living Philosophers, pp. 305-32.

Stoutland, F. [1985]. 'Davidson on Intentional Behaviour', in LePore, E. And McLaughlin, B. P. (eds.), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Blackwell, Oxford.

Taylor, C. [1964]. *The Explanation of Behaviour*. Routledge and Kegan Paul, London.

Taylor, R. [1966]. *Action and Purpose*. Prentice-Hall, Englewood Cliffs, N.J.

Tuomela, R. [1977]. *Human Action and its Explanation*. Dordrecht, Reidel.

von Wright, G. H. [1963]. *Norm and Action*. Routledge & Kegan Paul, London.

von Wright, G. H. [1971]. *Explanation and Understanding*. Cornell University Press, Ithaca.

von Wright, G. H. [1972]. 'On So Called Practical Inference', reprinted in von Wright, *Practical Reason. Philosophical Papers Vol. 1*. Blackwell, Oxford (1983), pp.18-34.

von Wright, G. H. [1974]. *Causality and Determinism.* Columbia University Press, New York.

von Wright, G. H. [1976a]. 'Replies', in Manninen, J. and Tuomela, R. (eds), *Essays on Explanation and Understanding.* Reidel, Dordecht and Boston, pp. 371-413.

von Wright, G. H. [1976b]. 'Determinism and the Study of Man', in Manninen, J. and Tuomela R. (eds.), *Essays on Explanation and Understanding.* Reidel, Dordecht and Boston, pp. 415-35.

von Wright, G. H. [1980]. *Freedom and Determination.* Acta Philosophica Fennica,vol. 31, No.1. North-Holland Publishing Company, Amsterdam.

von Wright, G. H. [1981]. 'Explanation and Understanding of Action', reprinted in von Wright, *Practical Reason. Philosophical Papers Vol. 1.* Blackwell, Oxford (1983), pp. 53-66.

von Wright, G. H. [1984]. 'A Reply to my Critics', in Schilpp (ed.), *The Philosophy of Georg Henrik von Wright*, The Library of Living Philosophers.