

FREE AGENCY  
A NON-REDUCTIONIST CAUSAL ACCOUNT\*

Wilhelm VOSENKUHL  
Universität München

The objective of this paper is a defence of the thesis that an action may be called “free” if it has rational causes. But it is neither clear what an appropriate causal account of action would be like nor what rational causes are. Are rational causes a special type of causes and if so, how do they relate to other causes of actions? We might suspect that we are forced into reductionism and ineluctable determinism anyway by a causal explanation of action. This would restrict any account of free agency if ‘free’ implies that the agent has the power to determine his own course of action. Self-determination would thus be ruled out by determinism. But if we find reasons to subscribe to a causal account of agency it would be counter-intuitive and self-defeating to ignore those causes which are not rational ones.

Talking about ‘causes’ of actions does not make things clearer as actions in general do not give any direct evidence about their causes at all. We do not even know which one of the many possible descriptions of an action we should adopt, if we don’t know anything about possible reasons that caused it. Was it a big squeeze, a strangulation, a friendly kiss or an attempt to restore life? We would certainly know enough to identify it, if we knew the agent’s reasons for doing it, provided he did it intentionally.

I

Let’s come back to our first question and ask, what are the causes of actions. Donald Davidson (1963) proposed: its *reasons* are its

\* I am indebted to Martin Hollis whose criticism helped me to face a number of problems some of which, I suspect, nevertheless remained unresolved.

causes, more precisely its primary reasons, consisting of a pro-attitude<sup>1</sup> toward actions of a kind and a related belief that the action he is performing is of that kind (cf. 685f.). Davidson claims that primary reasons rationalize actions and that rationalization is “a species of ordinary causal explanation” (685).

There are at least two major problems connected with this account of Davidson's. First, are reasons causes? And second, how can we explain action by an ordinary causal explanation? The first problem has been debated at length in the literature, so I should be allowed a short answer. The second problem is – I think – more interesting, so it might take longer to cope with that.

(1) Arguments against the explanation of reasons as causes have mainly been put forward by philosophers like Melden, Hamlyn, Peters and Winch, dubbed “philosophical psychologist” by Ruth Macklin (1969, 388-415). One of these arguments states, that reasons in terms of propositional attitudes cannot be causes from logical grounds, as beliefs, intentions, wants etc. are not events and therefore of no use in a causal account (cf. G. Ryle, 1965, 113). This argument is not sound, because it confounds the propositional content of beliefs, intentions and wants with the event when they are activated by a person. Most propositional attitudes are mental states and could not be looked upon as events, if they were not activated by an indexical belief. This indexical belief, which could be characterized this way: ‘I believe that something is actually the case here (there), now (at time  $t_1$ )’ is dependent on the agent's belief context and recalls some of his context-independent beliefs, intentions and wants. Indexical beliefs and context-independent propositional attitudes are the ingredients of mental events that are reasons and causes of actions<sup>2</sup>. For an indexical belief it is not only crucial that the agent comes to believe that something is the case in a certain place at a certain time, i.e. that he or she is able to locate his or her belief. It is equally important that the person is able to locate his or

1. Pro-attitudes include “desires, wantings, urges, promptings, and a great variety of moral views, aesthetic principles, economic prejudices, social conventions, and public and private goals and values in so far as these can be interpreted as attitudes of an agent directed towards actions of a certain kind” (686).

2. I would refrain from using an obscure phrase like “onslaught of a state or disposition” as D. Davidson does (1963, 694) to explain a mental event.

her belief by using explicitly or implicitly the “essential indexical” ‘I’. This indexical is essential – as John Perry (1979) pointed out – insofar as the replacement of it by other terms, even by the name of the agent himself, changes the meaning of the belief and destroys the explanatory force of the belief as an ingredient activating context-independent propositional attitudes<sup>3</sup>.

Another argument denying the reasons-causes-equation runs as follows: propositional attitudes are logically and not contingently connected with actions and therefore not their causes. A.I. Melden (1961, 53) argues in favour of the logical connection saying: “. . . nothing can be an act of volition that is not logically connected with that which is willed – the act of willing is intelligible only as the act of willing whatever it is that is willed”. So the explanation of wanting would be logically dependent on the wanted. This argument is basically sound. There is a logical connection between reasons and actions because the description of a reason to do something will always describe the intentional object, but – as Frederick Stoutland pointed out (1970, 124f.) – not a strong type of logical connection. There would be a strong type connection if reasons could be distinguished from each other by the effects they have. But this does not hold. The occurrence of an act of will is logically independent of its fulfillment. It is true that acts of will always have as part of their internal structure intentional objects. But intentional objects are neither effects nor fulfillments. So the relation between reasons in terms of mental acts and actions is a contingent one notwithstanding the weak logical connection between mental acts and their intentional objects.

There is a further argument which purports to show that reasons are not causes. It is derived from the contention that there are no general laws which would be required in a causal explanation of human action. General laws of this kind would presuppose the possibility of reducing mental events to physiological or physical events. And this is denied by the above mentioned ‘philosophical psychologists’ (cf. Melden, 1961, 93; Winch, 1958, 78; Peters, 1958, 148).

3. There are at least two major problems not to be discussed here which are involved in the notion of an indexical belief: firstly how an agent knows that it is he himself who ‘comes to believe something in a place at a certain time’, secondly how this belief is related to objects of belief, i. e. propositional contents. For both problems see John Perry (1979, 4ff.).

Their denial of general laws and reductionism is, of course, related to their denial of activated propositional attitudes being events. Although we argued against the latter we haven't yet given conclusive evidence about how we could identify mental events. Discussing the general law argument gives the opportunity to make up for this. This discussion leads over to the second problem mentioned above: how an explanation of action can be given on basis of an ordinary causal explanation.

(2) It should be clear that we subscribe to some kind of general law, which is valid for the relation between mental events and actions, if we maintain that actions are to be explained causally. But it is far from clear how we make use of the law in those explanations. Must we know the law explicitly or is it enough to presuppose some law underlying the singular causal statement 'person P caused action a'? And there is a further point to be made clear: Do we affirm reductionism in assuming some general causal law? If we follow Davidson the answer to the latter question is negative (cf. 1970, 87ff.) and with respect to the former he proposes that we do not need to know the law in giving singular causal statements (1967, 697ff.)<sup>4</sup>.

Indeed it would not be desirable to be forced into reductionism by a causal explanation, at least because we are aiming at a clarification of two kinds of causes and two kinds of determination in free action. Reductionism only allows for one kind of determination and thus does much to encourage the spectre of determinism. But our interest in elucidating two kinds of determination which make free action possible is only a motive not a reason for being opposed to reductionism.

It seems after all that the general law argument is not to be criticized for its anti-reductionism<sup>5</sup>. Still, a psychologist would maintain that avoiding reductionism does entail giving up causal explanation. If he did not so he would lose a strong argument against the causal approach.

The key to a non-reductionist causal account of actions is in the

4. I am not dealing with the question whether causes can be wholly particular (cf. Mackie 1965) as I do not see its immediate relevance for the problem of reductionism, and from reasons made clear further down.

5. I do not consider the difference between strong and weak reductionism here; cf. R. Macklin (1969, 396f.).

way we make use of causal laws in causal explanations. If we don't have to know those laws in our explanation – as Davidson claims – how do we make sure our explanation is a causal one? How do we justify our accepting a singular causal statement as being true, if we don't know the underlying causal law? Davidson argues:

“that we must distinguish firmly between causes and the features we hit on for describing them, and hence between the question whether a statement says truly that one event caused another and the further question whether the events are characterized in such a way that we can deduce, or otherwise infer, from laws or other causal lore, that the relation was causal”. (1967, 697).

Davidson pleads that we must consider what we really do in a causal explanation: we relate *statements* describing events not events themselves. So, whenever we ask for the truth of a singular causal statement we ascertain first of all, if the statement ‘*p* caused *q*’ relating two true descriptions of event *p* and *q* is a true one. Having proved the *truth* of the sentence *in a language* we could then ask, if from our description of the cause *p* together with some appropriate causal law an effect like *q* would follow. Thus there is a difference between an analysis of causal explanations and the causality of two events. Knowing that the singular causal statement ‘*p* caused *q*’ is true does not presuppose knowing the causal law covering *p* and *q*. But knowing the first gives us good reasons – as Davidson believes – to assume an appropriate causal law underlying ‘*p* caused *q*’ (cf. 1967, 701). We would know there is some law without knowing what it was like. It should be clear now – at least from a Davidsonian point of view – how we make use of causal laws in causal explanation: we need not know what the law is, but the truth of the singular causal statement we use to explain the relation between two events gives us reasons to believe there is a causal law underlying the singular causal statement.

This account of the possibility of singular causal statements without knowledge of causal laws differs from others, say e.g., the one J.L. Mackie (1965) offered, insofar as it doesn't explain causal relations in terms of conditionship relations. Following the difference he makes between descriptions of events and events themselves Davidson would not look upon singular causal statements as necessary or sufficient conditions. Those conditions are parts of causes but not parts of their descriptions. Davidson is therefore not only relieved of

giving a fully-fledged analysis of causes in terms of conditions, but he is also able to allow causes to be called necessary whenever they are described as sufficient (cf. 1967, 698). The reason for the latter is that there seems to be a certain symmetry between the description of causes and effects: the fuller the description of a cause, the better the chance that it is to be shown as sufficient and the worse as necessary; the fuller the description of an effect, the better the chance that the cause is to be shown as necessary, the worse as sufficient (loc. cit.).

One could doubt that this symmetry is of any help as soon as a full description of a cause shows that it is, what Mackie called, an “INUS-condition”, an “insufficient but necessary part of a condition which is itself unnecessary but sufficient for the result” (1965, 245). Imagine a person P addressing another one A for the first time asking what the time was. Let’s assume P’s pro-attitudes were that he always believed the addressee A was a nice and interesting person and that P had the intention to address A for a long time. But yesterday – these being P’s indexical beliefs – he caught sight of him and felt he should not wait any longer before getting into contact with him and that he could do this by asking what the time was. These indexical beliefs could be considered to be an insufficient but necessary part of P’s primary reasons to ask A what the time was; these primary reasons being themselves an unnecessary but sufficient condition for that question. Without doubt, P could have had other reasons as well to ask A for the time, although the ones he had were sufficient. And P’s indexical beliefs might not have been sufficient in themselves, though they were necessary insofar as P wouldn’t have activated his pro-attitudes without them. We certainly could analyse P’s indexical beliefs as an INUS-condition and his primary reasons as a ‘minimal sufficient condition’ (cf. Mackie, 1965, 246ff.) for P’s asking A what the time was.

But this analysis is besides the point. We are not interested in the causation of that very question but in the causation of an event we could describe as (E) ‘P’s long intended first addressing the highly esteemed person A yesterday’. P’s asking A what the time was is nothing but an attributive part of a fuller characterization of that first addressing A. If description (E) is an adequate description of P’s action and the above given primary reasons an adequate description of its cause, we must not jib at holding the latter to be described as

sufficient *and* necessary for P's action under description (E). There is no INUS-condition in the *description* of P's mental event which caused his addressing those very words ("What's the time, please?") to A. The mental event is *described* as a single and unique cause, and its description is not the *event* itself, i.e. in ontological terms. The description is the *analysans*, not the *analysandum*. The event as *analysandum* might well be characterized in terms of conditions. But we should not forget that giving characterizations of this kind implies – at least in a causal explanation of action – the problem of attaching ontological predicates to entities we have no empirical evidence of. These brief considerations should throw doubts upon an explanation of causal relations in human action in terms of conditionship relations.

So far we have argued that, first concerning our use of causal laws it is possible to give a causal explanation without knowing the law covering two events described in a singular causal statement and second, we aren't compelled to analyse the event described as the cause of the other in terms of conditions. Both results however still leave it open whether we can give a non-reductionist causal account of action. This problem is closely connected with the question how we are able to identify mental events like the one we described above. The identification of those events might commit us to some undesirable causal mechanism connecting these events to describable psychophysical events in our brain. We might thus incur the liabilities of a psychophysical reduction of the mental and not come to grips with that kind of mental determination we reckon crucial for free action. For doing the latter it is not enough to be able to give a causal explanation without knowing the covering law. We must be in the position to show, that the assumed law does not determine the agent's primary reasons, i.e. the mental event described as the cause of his action.

Again, Donald Davidson offered a possible solution to this problem by identifying mental events with physical events (1970, 82ff.). He called an event "physical, if it is describable in a purely physical vocabulary, mental if it is describable in mental terms" (*loc. cit.*). According to Davidson we have physical and mental open sentences where we fill in the description operator to pick out certain events. He adopts an identity theory between the mental and the physical which he dubbed "anomalous monism" (1970, 87). This monism im-

plies that not all events are mental but that all events are physical and it claims that mental phenomena aren't to be explained in a purely physical manner.

The decisive question concerning this kind of monism is: how does it guarantee the identity of a mental with a physical event without being reductionist? After all, there must be some kind of reduction involved in the identification of a mental event with a physical one, which is implied in the very act of identification. But it is not this identification that merits the name 'reductionism'. The latter would presuppose a psycho-physical law correlating mental events in terms of propositional attitudes and physical events in terms of neuro-physiological happenings. Therefore our question is: how is an identification of a mental event with a physical one guaranteed without its being based on any psychophysical law?

Davidson falls back on the supervenience-hypothesis to elucidate this alleged lawless identity, saying:

"there cannot be two events alike in all physical respects but differing in some mental respect, or. . . an object cannot alter in some mental respect without altering in some physical respect" (1970, 88).

We could analyse this hypothesis in the following way:  $\phi$ -concepts are supervenient to  $\psi$ -concepts, or  $\phi$ -concepts are determined by  $\psi$ -concepts, if and only if (i)  $\phi$ -concepts are not reducible to  $\psi$ -concepts and (ii) it is impossible for two events to be exactly alike in their  $\psi$ -characteristics but different in their  $\phi$ -characteristics.

Surely, if Peter and Paul were exactly alike in their physical characteristics they could not be different in their mental characteristics and would always e.g. understand each other's utterances. But this would not hold for Peter and Paula, at least not with respect to their understanding each other's utterances, which would at least presuppose that they would not have any physical differences. What does the supervenience-hypothesis prove according to the non-reducibility of mental and physical events? I think nothing. But it is an elucidation of what is meant by lawless identity. We could paraphrase this kind of identity by saying: even if we knew all of the truths about the physical characteristics of two persons at a time  $t$  we would not have all of the evidence necessary to know what they believe, intend, want etc. at  $t$ . The meaning of those physical characteristics is not to be equated with the meaning of their propositional attitudes. Physical meaning does not entail mental meaning. Even if

a translation manual in terms of psychophysical laws were possible, indeterminacy between the two kinds of meaning wouldn't be ruled out.

The point of indeterminacy is that we don't have identical extensions with physical and mental predicates. The reason why is that there is no predicate of syntax like the mental predicate 'true-in-L' in a purely physical language that is applicable to all and only the true sentence of that language (cf. Davidson, 1970, 88). So whenever we have an open sentence in physical terms like 'event  $x$  is  $\psi$ ' there can't be a psychophysical law that would have the form of a biconditional '( $x$ ) ( $x$  is true-in-L if and only if  $x$  is  $\psi$ )' where we substitute a physical predicate for  $\psi$ . We cannot generally reduce truth in such a language to a syntactical property (cf. Davidson, loc. cit.).

Now, we might admit indeterminacy for these or similar reasons and allow plausibility to the supervenience hypothesis but still not be prepared to accept a lawless identity of physical and mental events. For what points could there be in talking about the identity of two events whose predicates don't have the same extension? We should notice the direction of inference concerned in the identification business on the one side and in reductionism on the other. In reductionism we presuppose causal knowledge in terms of physical knowledge of an event and deduce its mental meaning. Anomalous monism claims that this is impossible. In the identification of mental events with physical ones we presuppose a knowledge of the meaning of the mental event and try to describe this event by some physical event that is coextensive with the mental one<sup>6</sup>. There is again no knowledge of a law involved but some inductive generalization correlating each mental event with a number of physical events. We can only proceed from the mental to the physical but not the other way around. This is what we mean when we deny reductionism but nevertheless subscribe to a law which covers a mental cause and an action. The law is a physical one, that holds between the physical event describing the mental event and the physical event describing the action. It neither relates the mental description of the cause with the action nor the mental event with its physical description.

Now it should be evident in what sense we are able to give a non-reductionist causal explanation of action, an explanation without

6. As to Davidson (1970, 90): "... there seems no compelling reasons to deny that there could be coextensive predicates, one mental and one physical."

knowing the underlying causal law. And it should be plain why we cannot explain causes of actions in terms of conditions. Those conditions would be essentially physical and could not replace the description of a mental event as a cause described as necessary and sufficient for an action under a certain description. Since these two points are evident we are in a position to retort to the above psychologist that we are able to deny reductionism and affirm a causal explanation at the same time.

## II

As we are able to give a non-reductionist causal account of action we have the means to distinguish between two kinds of determination: determination by a causal law and determination by mental events. This is the core of what is named anomalous monism by Donald Davidson and the essential feature of Kant's – at least implicit – theory of moral action based on the third antinomy in his 'Critique of Pure Reason'<sup>7</sup>. Kant would call determination by mental events 'autonomy', identifying an autonomous agent with a free agent. The subject of determination is either a causal law or a rational agent with both kinds of determination coexisting in every instance of action called 'free'.

I would like to couple some of Kant's ideas about 'free will' with the causal account of action outlined in the first section. The reason for doing so is that Kant – as far as I see – gave a causal account of free will which happily coincides with our account of action.

(1) One might see a problem in his identification of free action with moral action. But I think that this has often been misconceived by Kantian scholars as a constraint on free action confining it to moral action. This misunderstanding comes from the hardly disputable fact that Kant didn't offer a fully-fledged theory of action. Although I am not able to argue this point here in detail I would maintain that there is a general non-moral concept of free agency in terms of rational agency in Kant's philosophy based on the third antinomy, this being itself the basis for moral agency<sup>8</sup>. For if it were not pos-

7. This was noticed by John Yolton (1966) among others and by Davidson (1970, 79, 101) although both did not make use of it in an account of free action.

8. Cf. W. Vossenkuhl (1979) for more detailed arguments to this point.

sible for a rational agent to be self-determinant within nature, i.e. notwithstanding the causality of nature, it would not be possible for him to act morally. Being a self-determinant agent, i.e. having the rational power to start – as Kant said – a “causal chain”, is essential for free agency moral and non-moral.

Though I don't want to reconstruct Kant's idea of causality or his idea of freedom blow by blow here, it is interesting to note that he tackled the problem of a non-reductionist causal account of rational action by distinguishing between the causality of nature and the causality of freedom. Actions are parts of the causality of nature as far as they are empirical, spatiotemporal entities and thus determined by causal laws (cf. KrV, A 550, B 578). But they would not be perceived as actions of a rational subject solely from that kind of determination. To know an action was caused by a rational agent presupposes, at least for Kant, that it was caused independently from any causal law valid in nature. The epistemological point in this argument is developed in the 'Critique of Judgment' (cf. § 75) where Kant claims that it is an essential of human perception to understand the causality of nature in analogy to the causality of freedom. He formulated this analogy in his “subjective principle of reflecting judgment”. We could paraphrase his position by saying that the causality of freedom is a subjective transcendental precondition of the perception and explanation of the causality of nature without being its objective condition. Thus the causality of freedom is thought by Kant as a causal power both of action and of perception. This entails that it could neither be thought as possible for an agent endowed with reason to perceive nor to act independently from nature's determinism if there were not a causality of freedom.

The core of this account of what was called anomalous monism by Davidson is the concept of a rational agent whose distinctive feature is to intervene in the causality of nature. The decisive ingredient of this intervening-power is human reason. But how can reason be causal? Kant's answer is: by determining human will. If human will is not determined by contingent wants and intentions and thus free from any heterogeneous influence it is 'pure will', i.e. *rational* will. But where does this rational will get its goals from? Kant's answer to this is: from reason itself. To understand what is meant by this we should consider that his idea of reason implies a teleological aspect elucidated by the “idea of mankind”. This idea is a rational goal inso-

far as every person is an end in itself and not a means, neither for himself nor for others (GMS, 52). Whenever this being an end in itself is embodied in a person's action its result is autonomous, free action. To understand this pattern of causation one should not dwell on finding any material sense in the end-in-itself condition. Besides the teleological meaning of this condition in the "categorical imperative" the end-in-itself condition plays a methodological role in rational action. It restricts the means-ends relation involved in free action to one single type, namely causal one. Thus the type of causation is the same in nature and in free action with the only difference that in the latter reason in terms of pure will is a substitute for a natural cause<sup>9</sup>. The type of causation is identical although the causes are different. This is the background of Kant's explanation of the 'categorical imperative' as a "practical rule" and as a "product of reason" prescribing an action (KpV, A 36).

One might ask if there is any point in drawing upon this account of free action in our context of analysis. I think there is a point in it, because it reaffirms that the explanation of free action is dependent on the explanation of its causes. And it corroborates the intuition that free action is not a special kind of performing action but a special kind of causing them. What we can learn from Kant is what most of us would not have denied anyway, that free action is dependent on a special kind of reasons, namely *rational* ones. Having agreed with this the next two moves are not common at all: first to realize that it is special about free actions that their explanation amounts to their justification and second to emphasize that 'free' is first of all not the predicate of an action but of its cause, as it can only be the predicate of a true description of a contingent action, if that action had rational causes. If this is so, then there is no class of free actions but only a class of free in terms of rational causes. From our interpretation of Kant's account an action is described 'free', if the agent chooses it *as if* it was caused by 'pure will'. In short, *an action is free if it has rational causes*. There is only a sense in this tautology if we apply it to distinguish the causes of actions but not their effects.

So far we have answered the last two questions raised above: it is a special type of causes that makes actions 'free actions': a type we could name *rational causes*. This answer implies that there are other

9. Cf. Kant, Handschriftlicher Nachlaß, Reflexion 5612, Akad.—Ausg. vol. 18.

causes of actions as well that we could summarize as *natural causes*. Those causes would include all events inside and outside an agent that cause him to do something, like bumping into another's car if the driver was dazzled.

(2) Most of us still would not be satisfied by the account of rational causes just offered. The reason is that our normal use of 'rational' does not suggest a conceptual distinction between 'natural' and 'rational' like the one made above. The meaning of 'rational' in the context of our causal approach is not identical with the meaning of 'rational' e.g. in a teleological framework of practical reasoning. Rationality in practical reasoning is a relation between means and ends, a relation between certain goals and appropriate actions. If these actions are appropriate to bring about their goals in an optimal way they are rational and it is rational to follow certain goals if we have the means to make them true. Rationality in practical reasoning is thus dependent on empirical, i.e. contingent facts and assumptions concerning the quality of actions as appropriate means for certain ends and the quality of ends in terms of their achievability. Rationality understood in this way is based on the agent's experience and his calculative power to choose the best possible action under circumstances.

It is the dependence on these contingent features of a teleological concept of rationality that we cannot make use of it in a causal account of free action. What we need is an *a priori* concept of rationality as we are not prepared to allow natural causes being involved in free agency. This envisaged non-empirical, *a priori* concept of rationality cannot stand for a relation between propositional attitudes and the empirical. It can only be used for the relation between propositional attitudes themselves<sup>10</sup>; it is *a priori* insofar as it does not qualify the content of beliefs, wants, intentions etc. but their *coherence*. Rationality understood in this way does not denote a calculative power to make choices. A definition of the rationality of propositional attitudes in terms of their coherence gains from the definition of a rational belief offered by Martin Hollis. He defined a belief *p* as *rational* "if and only if there is a belief *q* such that *q* supplies a reason for holding *p* and *p* does not entail the falsity of *q*" (1970, 235). Martin Hollis holds that rational beliefs must have a structure

10. The concept of an *a priori* rationality of beliefs was defended by M. Hollis (1970).

which can be there even when some of them are false. His definition of rationality can be applied to determine the coherence of propositional attitudes in general. *Propositional attitudes are rational if and only if each set of the set of beliefs, wants, intentions etc. is coherent and those sets are coherent with each other.*

Let us elucidate this definition with inconsistent propositional attitudes. Following the idea of coherence it is rational to believe *that p* and *that q* where *q* implies not-*p*, if one sincerely entertains a belief implying that there is a reason *r* for *p* and *q* and that the inconsistency between them does not have intolerable consequences. One might believe that some object has a curing power and that this power did not work in a certain case. This inconsistency is tolerable, if the same person sincerely believes that one of the essential preconditions for an actual curing effect did not apply. For such a case one would find more restrictions desirable than *sincerety* alone. If there were direct evidence to the opposite, say, that all preconditions applied without any curing effect caused by our dubious object, *q* would entail the falsity of  $r^{11}$ . It would not be rational to believe in its curing power. The intolerable consequences of our inconsistency would be that there is no curing power in that object and that both inconsistent beliefs would be pointless and void.

Could we allow of any as yet unknown precondition saving our belief in the object's capacities? If we were able to confirm that this precondition is not really a new one but always applied without being noticed, we could certainly admit of the new interpretation. But the rationality of our belief could not suffer any definitely new precondition. If the curing-procedures of our object are not limited but open to indefinitely many amendments consisting in new preconditions and not just new interpretations of old ones, our belief in the object's curing power would not be a rational one. We either realize that we need a set of beliefs ruling out intolerable consequences of some inconsistency or we are confident that we will always find reasons explaining why those consequences never occur. This confidence however is not rational as it is not coupled with the belief that there actually is an inconsistency with potentially intolerable consequences at all. Rational confidence is based on an available,

11. *p* would entail the falsity of *r* if our object always proved its curing capacity; in this trivial case *r* would be wrong because *q* would be wrong in the first place.

limited set of beliefs making some inconsistency tolerable. The confidence criticized as irrational comes down to the plain contradiction that there is an inconsistency and that there is no inconsistency at the same time. Either there is an inconsistency that needs explanation to be tolerable or there is none. In this case there is no reason for any confidence in possible explanations as those explanations are not impossible but pointless.

Two inconsistent *beliefs*  $p$  and  $q$  are *rational* in terms of coherence

- (i) if they are sincerely held and
- (ii) if there is a limited set of beliefs, at least one belief implying that there is at least one reason  $r$  for  $p$  and  $q$  and neither  $p$  nor  $q$  entails the falsity of  $r$  and the additional belief
- (iii) that there are no intolerable consequences following from holding  $p$  and  $q$ .

We could argue along similar lines to expound the coherence of beliefs and wants. A person's belief *that*  $p$  and his or her want *that*  $w$  are coherent, if he or she further believes, that the want's fulfillment  $w$  is consistent with belief  $p$ . But what if one wants that  $w$ , say that he will fly to the moon, and knows he cannot afford it? If this person believes that he will only be able to realize those wants he can afford, there should be no problem with the coherence of the person's beliefs and wants. But wants can be terribly strong although one knows they could or even should not be accomplished or enacted. Executing a want against one's conviction that it is useless, void or even wrong to do so, is certainly inconsistent and not rational. But what if one is convinced that it is good for him to follow his wants and performs  $w$  instead of recognizing his own better judgements? This person, who could be called incontinent, is not suffering from a conflict between his beliefs and wants but rather from an inability to decide about colliding beliefs. Some of those beliefs are affirmed while others are seen but nevertheless remain unacknowledged without any reason. This man does not hold – as Donald Davidson pointed out (1969, 112) – logically contradictory beliefs but “acts, and judges, irrationally, for this is surely what we must say of a man who goes against his own best judgement” (loc. cit.). He acts irrationally as he has no reason to ignore his own best judgement. His beliefs are not coherent because he neither holds a belief implying at least one reason for his ignoring his own best judgement nor the additional belief that there are no intolerable consequences following from his

conflicting propositional attitudes. As this man is ignorant about the beliefs we claim to be crucial for the coherence of beliefs, his actions are not determined by rational causes.

Now imagine our man comes to believe that there are intolerable consequences following from his actions. He nevertheless holds (*that p*) that it would be better for him to realize his want *that w* but refrains from its fulfilling because he is threatened by sanctions impending upon *w*-actions. He believes that there are sanctions awaiting his actions and therefore does not intend *i* to fulfill his want. Would this belief not count as a sufficient reason  $r_1$  for not intending *w* although *p* and *w* are coherent? This belief would imply the additional belief that there would not be any intolerable consequences if he abstains from *w*. Would we not be obliged to concede that our man's conduct is rational?

His intention not to fulfill his want is certainly 'zweckrational' but this does not make his propositional attitudes rational in terms of coherence. Firstly our man's want *that w* and his want (*that w<sub>1</sub>*) not to suffer any sanctions for fulfilling *w*, i. e. his fear of sanctions, are conflicting. This means that his wants *w* and *w<sub>1</sub>* are incoherent. So are his beliefs that he would better realize *w* (*that p*) and that he would suffer intolerable consequences for *w* from sanctions threatening him (*that q*). Secondly although there is a reason  $r_1$  not to intend *w*, this reason's falsity is entailed by *p* and *w*. Presupposing the coherent propositional attitudes *p* and *q* intention *i*, not to act according to *p* and *w*, is not rational. This intention relates to *q* and *w<sub>1</sub>* and is coherent with these propositional attitudes. But in connection with all beliefs (*p*, *q*) and wants (*w*, *w<sub>1</sub>*) intention *i* is incoherent. As *p* is incoherent with *q*, and *w* is incoherent with *w<sub>1</sub>*, intention *i* cannot be coherent itself.

Notice that even if certain beliefs are coherent with certain wants the resulting intentions are incoherent if neither the set of beliefs nor the set of wants are coherent themselves. Thus the rationality of intentions is not only dependent on the coherence between beliefs and wants but on the coherence amongst both beliefs and wants, independent of each other.

We could weaken this last condition for the rationality of intentions by demanding coherence of propositional attitudes for all beliefs, and for the relation between all beliefs and some wants, but not for all wants. If the man in our example would not believe *that p* but

rather affirm his own best judgement that it is not legitimate to accomplish  $w$  he would have good reasons to deny himself his want  $w$  and coherently intend  $i$ , not to realize  $w$ . In this case he would not refrain from fulfilling  $w$  just as he feels threatened by sanctions ( $r_1^*$ ) but on basis of a reason  $r_1$  which is not falsified by one of his beliefs. He would thus practice temperance instead of incontinence.

This final modification of our agent's propositional attitudes not only meets the fact that human agents are not able to choose their wants but also a realist account of human wants. One of the features of a realist account of human wants is that they cannot be 'rational' independent from beliefs. Whenever some want is rational it is primarily because of its coherence with some belief but not because of its coherence with all other wants. Whatever wants and wishes human agents entertain and have to cope with, there is no point in demanding that they should be all coherent with each other and with all of those agents' beliefs, if his intentions' rationality is preserved.<sup>12</sup> An agent's *intentions* are *rational*

- (i) if his beliefs are coherent,
- (ii) if he has a reason  $r_1$  to abstain from those wants which are incompatible with his beliefs,
- (iii) if this reason is not falsified by one of his beliefs,
- (iv) if all other wants are coherent with his beliefs.

Provided an agent's beliefs, wants and intentions are coherent the way we proposed then his actions have *rational causes*. Coherent propositional attitudes are the causes of free actions. From the way we explained the causation of action above those propositional attitudes must be activated by an agent's indexical belief held in some spatio-temporal context. These activated propositional attitudes together with the indexical belief will then rationally determine his action.

(3) We have distinguished two concepts of rationality to be able to say that rational causes of free actions are *coherent propositional attitudes*. This explanation still needs some amplification as we might be suspicious about the way those causes really work in action. We have another distinction at hand that runs parallel with the two kinds of rationality and which shows how coherent propositional attitudes cause actions. This distinction is the one between determination and

12. The same applies to other non-cognitive, optational propositional attitudes like fears and fancies, needs and desires etc.

choice. To say that coherent propositional attitudes are rational causes of actions means that these causes *determine* those actions. If a person can be characterized by coherent propositional attitudes and if they are the causes that determine his actions those causes can't be chosen.<sup>13</sup> The reason why is explained by Thomas Nagel: a person does not choose those causes, "for choices must issue from him if they are to be his, and this means that they must be the product of determining principles which constitute him as the source of his choice, and which could not be chosen by him because in their absence there would be no *he* to choose" (1970, 23). The role of Nagel's "determining principles" in his moral analysis of altruism is played by the coherent propositional attitudes in our non-moral analysis of free action. A person cannot choose the coherence of his propositional attitudes as he can't choose the wants and intentions that characterize him as the very person he is. This implies that a person can't choose to act from rational causes. A free agent is defined by the coherence of his propositional attitudes in the same way as a moral agent is defined by the principles that constitute him.

With respect to Kant's identification of free and moral action we now roughly see that we are able to modify his approach without changing his moral theory. A decisive point in a more detailed analysis of this thesis would be an argument for the interpretation of Kant's concept of 'pure will' with our concept of the coherence of propositional attitudes. If this argument is possible then the explanation of moral action presupposes the explanation of free action, as the coherence of propositional attitudes is the precondition of a person's being determined by moral principles. This means that free action must not be identified with moral action. An argument like this might be the first step into the direction of reconciling a theory of action with a Kantian theory of morality.

Having explained what rational causes mean and how they work in the causation of free action we are able to sketch the relationship between the coherence of propositional attitudes and the rationality of choice. The two types of rationality seem to be compatible under the condition that the rationality of choice conforms to the agent's beliefs, wants and intentions. If an agent were to choose an action solely

13. The same applies, although it is not argued here to indexical beliefs. I cannot choose that 'I' (or rather somebody else) come to believe something in a belief context.

on basis of a means-ends calculation and thus optimize the effect of his acting without making sure that the reasons for his choice conform to the propositional attitudes that characterize him as the person he is, his action would neither be free nor rational. The reason is not just a lack of reasoning about the conformity between his propositional attitudes and his principles of choice. If those principles do not conform to his propositional attitudes he would not be a person with coherent propositional attitudes. He would want something incoherent with his beliefs or other wants and intentions. It only appears now that both types of rationality must conform to allow a certain action caused by coherent propositional attitudes to be a rational one. We should now have enough reasons to believe that rational causes have to be presupposed for any kind of action that we are able to explain and justify as free action.

#### *Literature*

- Davidson, D. (1963): "Actions, Reasons, and Causes", in: *Journal of Philosophy* 60, 685 – 700.
- (1967): "Causal Relations", in: *Journal of Philosophy* 64, 691 – 703.
- (1969): "How is Weakness of the Will Possible?", in: J. Feinberg (ed.): *Moral Concepts*, Oxford, 93 – 113.
- (1970): "Mental Events", in: L. Foster e.a. (ed.): *Experience and Theory*, Cambridge/Mass. 79 – 101.
- (1973): "Freedom to act", in: T. Honderich (ed.): *Essays on Freedom of Action*, London/Boston, 139 – 156.
- Hollis, M. (1970): "Reason and Ritual", in: B. Wilson: *Rationality*, Oxford, 221 – 239.
- Kant, I.: *Kritik der reinen Vernunft* (abd.: KrV, quoted after Akad.-Ausg.).
- : *Grundlegung zur Metaphysik der Sitten* (abd.: GMS), ed. by K. Vorländer, Hamburg <sup>3</sup> 1965.
- : *Kritik der praktischen Vernunft* (abd.: KpV, quoted after Akad.-Ausg.).
- Kenny, A. (1975): *Will, Freedom and Power*, Oxford.
- Mackie, J. L. (1965): "Causes and Conditions", in: *American Philosophical Quarterly* 2, 245 – 264.
- Macklin, R. (1969): "Explanation and Action; Recent Issues and Controversies", in: *Synthese* 20, 388 – 415.
- Melden, A. I. (1961): *Free Action*, London.
- Nagel, T. (1970): *The Possibility of Altruism*, Princeton.
- Perry, J. (1979): "The Problem of the Essential Indexical", in: *Noûs* 13, 3 – 21.

- Peters, R. S. (1958): *The Coucept of Motivation*, London.
- Ryle, G. (1965): *The Concept of Mind*, New York.
- Stoutland, F. (1970): "The Logical Connection Argument", in: *American Philosophical Quarterly*, Monograph Series 4, 117 – 129.
- Vossenkuhl, W. (1979): "Freiheit zu handeln, Analytische und transzendente Argumente für eine kausale Handlungstheorie", in: H. M. Baumgartner (Ed.): *Prinzip Freiheit*, Freiburg/München, 97 – 138.
- Winch, P. (1958): *The Idea of a Social Science*, London.
- Yolton, J. W. (1966): "Agent Causality", in: *American Philosophical Quarterly* 3, 14 – 26.

# The MONIST

*An International Quarterly Journal  
of General Philosophical Inquiry*

Founded 1888 by Edward C. Hegeler

Editor, Eugene FREEMAN

*Editorial Board:* William P. Alston, Monroe C. Beardsley, Lewis White Beck, William A. Earle, Dagfinn Føllesdal, William Frankena, Maurice Mandelbaum, R. Barcan Marcus, Richard Martin, Mary Mothersill, Joseph Owens, Richard Rorty, J. B. Schneewind, Wilfrid Sellars, John E. Smith, Richard Wasserstrom.

*Managing Editor:* Sherwood J. B. SUGDEN

*EACH ISSUE IS LIMITED TO ARTICLES ON A SINGLE GENERAL TOPIC. Communicate in advance with the Editor for special instructions. Two nonreturnable copies (one copy outside the U.S. and Canada) should be submitted NINE MONTHS PRIOR to the announced publication date. Papers should be prepared for anonymous submission to readers, and should be from 10 to 20 double-spaced typewritten pages in length, including notes.*

## SCHEDULED PUBLICATION DATES:

- |                |            |  |
|----------------|------------|--|
| Vol. 65, No. 1 | Jan., 1982 | Contemporary Continental Analytic Philosophy |
| Vol. 65, No. 3 | July, 1982 | The Philosophy of Wilfrid Sellars            |
| Vol. 65, No. 4 | Oct., 1982 | Metaphysics and Logic                        |
| Vol. 66, No. 1 | Jan., 1983 | Right Reason in Western Ethics               |
| Vol. 66, No. 2 | Apr., 1983 | Integrity and Purity in Art and Morals       |
| Vol. 66, No. 3 | July, 1983 | Categories                                   |
| Vol. 66, No. 4 | Oct., 1983 | The Basis of Political Authority             |
| Vol. 67, No. 1 | Jan., 1984 | The Philosophy of Mathematics                |
| Vol. 67, No. 2 | Apr., 1984 | Sociobiology and Philosophy                  |

**Editorial Office: Box 1908, Los Gatos, California 95030**

**Business Office: Box 600, La Salle, Illinois 61301**

**SUBSCRIPTION RATES:** United States: Annual (4 issues): Institutions, \$24.00; Individuals, \$15.00. Single copies: \$6.00. Foreign postage: Add 50 cents to single copy rate or \$1.00 to subscription.