Guido Löhrer

ACTIONS, REASON EXPLANATIONS, AND VALUES

Abstract
According to the orthodox causalist view, intentional action is a behavior which is caused in the right way by the reasons for which the agent acted and it is explained in terms of their causally effective reasons. Thus, reasons are considered to be causes, and common-sense action explanation in terms of reasons counts as a species of causal explanation. This paper will challenge the orthodox view with an array of objections that together support a teleological account interpreted as a value-based account. Actions realize values and unveil them. A reason explanation of a rational agent's goal-directed behavior invokes the value toward which the conduct was directed.

1. *Introduction*

There is a fundamental gap between things that merely occur or happen to us and actions we perform intentionally and for a reason. These kinds of behavior call for different explanations. According to the causalist theory of mind and action, both kinds of behavior are to be explained causally but they differ in their causal histories. Cautious causalists consider intentional action as a behavior caused in the right way (i) by the agent's reasons, or (ii) by a set of appropriate mental states of the agent, or (iii) by their neural correlates, or (iv) by a fact about something the agent believed, desired, or intended (see Mele 2013, p. 168; 2003, pp. 39, 52; cf. Goldman 1970 and J. Hornsby 1980). Intentional action, then, is to be explained by invoking reasons, mental states, or facts which cause the action and do it in the right way. Starting from Donald Davidson's *Actions, Reasons, and Causes* (1963/1980), the causal theory of action explanation has become the standard view. Nevertheless, it turned out to be utterly challenging to explicate the crucial normative expression "in the right way" in purely causal terms (see Davidson 1973/1980, pp. 78-79; 1976/1980, p. 267; 1978/1980, p. 87 including fn. 3; Mele 2003, p. 52). Thus, the causal theory of action explanation still faces serious objections.

On the one hand these objections draw on counterexamples citing so-called deviant causal chains which connect mental states with the caused behavior without providing a reason for the behavior in question. Here is a general pattern for the interesting part of these examples: An agent's intention throws her into a nervous state, whereby she entirely loses control over her behavior. Even so, as it happens her intention is realized in a wayward way but just as it was intended. In cases of this kind, though the explanation is causal, it is no action explanation because it explains an event which is no

action at all[1]. On the other hand it has been disputed that a causal theory succeeds in ascertaining those mental states of an agent which actually have become causally effective. Even when the causal chain runs in the right way, the account does not quite achieve what we expect it to do.

The persistence of the problems faced by the orthodox view suggests that we need to look for alternative explanations. One of the alternatives that have been put forward is a teleological account. According to this view the best explanation of an action invokes the function or purpose of the behavior or the goal that the agent directed her behavior at, rather than its causal history.

I want to show, firstly, that the strongest of these explanations should be regarded as a value-based action explanation and, secondly, that a value-based explanation really achieves what a causal action explanation claims to accomplish but fails to do. In what follows I will (2.) give a sketch of the orthodox view and (3.) its main difficulties. Then, I will (4.) contrast this view with the teleological theory which is interpreted as a value-based account and finally (5.) confronted with three objections. Actions realize values and they unveil them. A reason explanation of a rational agent's goal-directed behavior invokes the value which was supposed to be realized by the conduct in question[2].

2. *Causal Action Explanations – The Standard View*

An intentional action is what an agent does for a reason. Invoking this reason explains her action. This is the largely uncontroversial basis of action-theoretical considerations. In *Actions, Reasons, and Causes*, the founding document of the causalist theory of action explanation, Davidson asks for the relation between an intentional action and the reason for which the agent performed the action[3]. His account, which has become the orthodox view, brings two theses into play: the *Belief Desire Thesis* and the *Causal Thesis*. The first of them concerns the form of explanatory reasons, the second relates to the species of explanation.

In everyday life, reasons for acting are commonly given as follows: I am bundling waste paper because it is Wednesday. Bob did a headstand to impress Sally. But according to Davidson, reasons explain single actions only if they have the form of so-called *primary reasons* or if they are reducible to this form. Primary reasons – this is the first thesis – are composed of two mental attitudes, namely an agent's pro attitude toward a type of actions with a certain property and the agent's belief that the single action in question is an instantiation of this action type and displays this property (cf. Davidson 1963/1980, p. 5): Bob has a pro attitude toward actions that impress Sally, and he believes that doing a headstand here and now is an action that impresses Sally. Therefore he does a headstand.

---

[1] Mele 2003, p. 53: "No plausible version of *CA* [the causal approach; G.L.] identifies action with *non-actional* events caused in the right way".

[2] The subsequent discussion relies on preliminary work in Löhrer 2006; 2008; Löhrer and Sehon 2016.

[3] Davidson 1963, p. 3: "What is the relation between a reason and an action when the reason explains the action by giving the agent's reason for doing what he did?".

The phrasing is quite heavy-handed, and the expression "pro attitude" could hardly prevail in action theory. Instead it is now said that a primary reason is a belief-desire pair consisting of a desire to achieve a goal and the belief that the action to be explained is an apt means to achieve this goal.

(BD)   A reason which explains an action ϕ consists of an agent's desire for A and her belief that ϕ will achieve A.

Taken by themselves, neither the desire nor the belief explains the action. Common-sense action explanations may usually suppress either the one or the other component of the reason or they tacitly presuppose them. Why are you going to the café? I am going to the café because I want to meet my friends. Why are you going to the café? Because I believe that by going to the café I will meet my friends. However, according to the orthodox view, these reasons must be translatable into the form of primary reasons. This is the belief desire thesis, which concerns the form of explanatory reasons.

In a weak sense any explanation which relies on reasons of this form might be considered as value-based, albeit in two different regards. Firstly, my desire to meet my friends reveals that I am caring for a meeting with them and that I attach value to it. Secondly, with my belief that going to the café is an apt means to realize my desire I appreciate the instrumental value of such a behavior[4].

If a reason has the correct epistemic form of a primary reason, it satisfies only a necessary condition, not a sufficient condition of an orthodox reason explanation of action. Davidson's question for the relation between an intentional action and the reason for which the agent performed the action is not yet answered. Right here the causal thesis comes into play.

(C)   The explanatory primary reason for an action is its cause (cf. Davidson 1963/1980, pp. 4 and 12).

Connecting the belief-desire thesis with the causal thesis we obtain:

The agent's desire for A and her belief that action ϕ accomplishes desire A caused the agent's action ϕ. Thus, the action is explained only by that reasonable reason (the pair of mental attitudes or its neural correlate) which has become causally effective[5].

Davidson stresses that action explanation is a species of causal explanation. Reason explanation is a species of the genus causal explanation. Explanatory reasons for action form a species of the genus cause. They are those reasons that cause actions in the right way[6].

---

[4] Cavell 2004, p. xvi: "[B]eliefs and desires are inextricable from evaluations".

[5] Davidson 1974, p. 233: "Two ideas are built into the concept of acting on a reason […]: the idea of cause and the idea of rationality. A reason is a rational cause. One way rationality is built in is transparent: the cause must be a belief and a desire in light of which the action is reasonable".

[6] Cf. Davidson 1963, pp. 10-11; 1987, p. 109. Davidson (1984), p. 26: "[B]elief and desire seem equally to be causal conditions of action". Cf. Mele 2013, pp. 161, 164, 168.

3. *The Davidsonian Challenge and the Troubles with the Standard View*

It is important to emphasize that in *Actions, Reasons, and Causes* the causal thesis does not primarily serve pursuing a naturalization program. It rather involves the solution to an explanatory problem, for "a person can have a reason for an action, and perform the action, and yet this reason not be the reason why he did it" (Davidson 1963/1980, p. 9). So it could have been that I went to the café because I wanted to meet my friends and I believed that I could achieve this goal in this way. But it might also have been that I went to the café because I wanted to have coffee and I believed that I could achieve this goal in this way. Both reasons would justify my action. But if I go to the café only for one of these reasons, this reason alone explains my action. So there are Davidsonian cases where an agent has a couple of reasons that would justify her action though she acted only for one of them. Here is an example from Alfred Mele of what might be called a Davidsonian case:

> "Two different things, *T1* and *T2*, independently dispose Al to mow his lawn this morning. *T1* has to do with schedule-related convenience and *T2* with vengeance. Al wants to mow his lawn this week, and he believes that this morning is a convenient time, given his schedule for the week. But he also wants to repay his neighbour for the rude awakening he suffered recently when she turned on her mower at the crack of dawn, and he believes that mowing his lawn this morning would constitute suitable repayment. As it happens, Al's purpose in mowing his lawn this morning accords with one or the other of *T1* and *T2* but not both" (Mele 2013, p. 170).

With situations like this in mind, the causalists give to their arguments in favor of the causal account the form of a challenge to the non-causalist one[7]. For instance, in virtue of what is it true that Al mowed his lawn for the one reason and not the other, and that the one and not the other explains Al's action? The Davidsonian challenger then hastens to affirm that the causalist does not have this problem. According to the orthodox view the reason which causes the action in the right way is the reason which explains the action. Since action explanations are hybrid, i.e., both reason-based explanations and causal, and primary reasons for actions are their causes, the cause of the action picks the explanatory reason out of the set of justifying reasons. How is that possible?

Davidson distinguishes a metaphysical level of extensionally construed entities however described, i.e., the events, from the levels of their descriptions, e.g., their description as actions. Descriptions of events differ in their vocabulary. Any extensionally construed event, no matter how described, allows for a description phrased in physical terms which makes the event a physical event. Analogously, an event is mental if it can also be described in a mental vocabulary.

Causality is a relation between individual events on the metaphysical level, i.e., a relation of tokens of events, no matter how described. Talk of causality is justified only if

---

[7] In Löhrer and Sehon 2016, pp. 86-87, Sehon and I argue that the challenge is actually intended to be an argument for the causal theory of action.

the event tokens fall under event types which are linked such that causal laws can be phrased as true statements of laws in the form of universally quantified conditionals.

Singular causal statements, such as "event p caused event q", will explain an event q only if the link constitutes the exemplification of a law. Davidson suggests that causal laws can only be couched in a purely physical vocabulary. Thus, the thesis that the reason for an action is its cause has to be understood as follows. If there is a true mental description of an event, then there is also a corresponding true physical description of a causal relation which exemplifies a law. Therefore, as Davidson says in *Mental Events*,

"it is possible to know that a mental event is identical with some physical events without knowing which one (in the sense of being able to give a unique physical description that brings it under a relevant law)" (Davidson 1970a, p. 224).

So construed, the standard view does not achieve what it claims to achieve. It only assures that there is a physical cause corresponding to the explanatory reason, but it is not in a position to single out the explanatory reason from a set of possible justifying reasons. With this, the promise of *Davidson's Challenge* has gone. It is unquestionable that the intentional behavior of an agent has causes. When an intentional behavior is caused, the agent's intentions or their neural correlates play a causal role in causing this behavior. Nobody would deny this. But the standard view does not merely assert that there are causes of human behavior, it also claims that the causes of human action are *reasons* of the agent (cf. Löhrer and Sehon 2016, p. 92). Though, in establishing the correct reason for which an agent performed an action, assigning a causal role merely to primary reasons in general instead of a specific reason is no step forward.

The causalist might reply that – in spite of an abundance of examples – it is not the task of a causal theory of action explanation to explain individual actions. What is at stake is rather an ontological anchoring of the explanatory reason and, according to the causalist, non-causal explanations are lacking such a fact of the matter (see, e.g., Child 1994, p. 96). Therefore, the lesson to be learned from the causalist would be solely a matter of principle. Firstly, the explanatory reason has to be conceived as a reason the agent had before she started to perform the action. Otherwise the reason could not be the cause which brought about the action. Secondly, given the belief-desire thesis, one must look for the explanatory reasons among the mental antecedents of an action. Thus, explaining an action amounts to invoking its causally effective reasons, i.e. explaining an action calls for some kind of mind reading as it were.

Davidson seems to have conceived of his challenge as an inference to the best explanation. If it is true that there are reasons for acting which are the cause of the action, the sought-after connection between action and reason is found as well. In this case the causal fact makes it true that one reason rather than the other explains the behavior. However, with an inference to the best explanation, there is a quick answer available to non-causalists. According to the non-causalist, there are non-causal, e.g. teleological truth-makers which make non-causal or teleological action explanations true. Such a teleological truth-maker is the underlying teleological fact that the agent has directed her behavior toward the particular end in question. With that we are facing a stalemate. If there is no further independent evidence, none of both parties may claim to

possess an argument against the other and in favor of its own theory of explanation, at least not without begging the question. If Davidson's Challenge poses a problem to the non-causalist, it equally poses a problem to the causalist[8].

Let us suppose that Davidsonians can answer Davidson's Challenge, then there is a further ponderous objection left at any rate, viz. the argument from primary deviant causal chains. This argument is supposed to show that the theory of causal action explanation is inappropriate because it is too liberal. It says that even if we concede that reasons are causes preceding the action, the standard view supplies us with the wrong kind of explanations for events which are no actions at all.

Davidson's own example concerns a mountaineer who desires to get rid of the second man on the rope and believes that she could achieve this goal by loosening her grip on the rope. This thought unnerves her and causes her to involuntarily loosen her grip (cf. Davidson 1973, p. 79). Again, instantiations of primary deviance always conform to the same pattern. An agent adopts an intention to act. However, the intention does not cause the action but prompts a nervous state which provokes an involuntary bodily movement that accidentally realizes the intention. Thus, an agent's beliefs and desires cause a bodily movement which makes for the intended goal. All the conditions of the standard view seem to be met, and yet we are not inclined to admit that the agent's beliefs and desires which cause her behavior fit in with a reason explanation of this behavior.

Adherents of the causal account have tried to counter the problem in a causal-theoretical manner. The subtlest of these approaches is presumably Al Mele's. Mele (2003, pp. 54-58) conceives a myth about Promethean agents whose implanted neurophysiological mechanism precludes that nervousness becomes the cause of a behavior. The cortex of Promethean agents distinguishes command signals prompted by an instantaneously effective intention from those which are caused by an intermediary nervous state. Only command signals prompted by an instantly effective intention are transformed into executive signals to specific muscles. Since these executive signals are not produced by nervous states, no correlating bodily movements occur. Thus intentional action cannot be confounded with events which, though caused by the same primary reasons and having the same results, are no intentional actions at all.

However, this proposed solution faces an objection too. The argument from deviant causal chains charges causalism of accepting too many events as actions, while Mele's concept of action seems to turn out too narrow. Normally we do not dispute that actions can be accompanied with nervousness. In competition or during an exam it may even happen that a maximum capacity could not be achieved without the excitement prompted by the intention to achieve maximum capacity. Intuitively there is no reason to deny that in a situation like those the agent performed an intentional action (cf. Wilson 1989, p. 252; Sehon 1997, p. 202; 2005, p. 98; Löhrer 2006, p. 792). Notice that Davidson, of all people, considers futile all efforts to solve the problem of causal deviance with a causal apparatus. Further, "in the right way" is a normative expression which doesn't have its place in scientific causal explanation (cf. Schueler 2001, pp. 254-257 and pp. 259-261; Keil 2007, p. 75). Certainly this already holds true for the talk of

---

[8] For a detailed argument cf. Löhrer/Sehon 2016, pp. 89-93.

causal deviance. Causal relations obtain or do not obtain. Speaking of a deviance of causal chains seems to be reasonable only if one thinks of a deviation relative to an intended or expected course of events which has been found correct. The interpreter will notice it only *ex post eventu* (see Searle 1983, p. 139).

4. *Teleological Action Explanation*

The term "purpose" (Ancient Greek: *telos*) is used with different meanings. Firstly, "purpose" denotes the goal of an action which the agent sets herself. Secondly, the term "purpose" signifies the function which a thing, activity, or person fulfills, e.g., a heart has the function to pump blood; a shopping list is supposed to function as a memory aid. Regarding the second meaning, we have to distinguish between functions which belong to things, persons and behaviors in a natural, non-intentional way on the one hand, and functions which things, persons and behaviors have only because an agent has a purpose for them on the other. The same holds for teleological action explanations. When we think of purposes as natural functions, we normally deal with the study of action types. However, if we think of purposes as functions in an intentional sense, then action explanations aim at ascertaining the function which the agent has for her individual action and the goal that she has directed this action at.

(1) Ruth Millikan's teleo-functionalism is a teleological account of the first kind. In *White Queen Psychology and Other Essays for Alice* (1993) she gives a naturalistic reconstruction of intentional action in terms of proper functions or teleo-functions, i.e. the purposes a certain behavior has been selected for (cf. Millikan 1993; see also Sehon 2013, p. 737). Intentional and rational behavior is interpreted as an evolutionary pattern of adaption that, like animal flight behavior, includes a large spectrum of partial functions which might occasionally be called up without committing the individual to a certain pattern, because these mechanisms are always available in a redundant number and they can compensate each other if one of them fails. Millikan emphasizes that these teleo-functional mechanisms do not have to occur in the life of every member of a species. It will do if they are sometimes useful in the evolution of the species and can be reactivated when required. It may be that rational behavior can be evolutionarily explained as a survival benefit, but presumably there are neither strict laws which make for predicting human behavior nor ceteris paribus laws of the behavior of those who have a certain teleo-functional property.

With the help of a theory of natural selection, a teleological account of this kind reduces the goal-directedness of a human behavior to its survival benefit. Yet a couple of obstacles would have to be removed in order to explain actions that seem irrelevant toward differential reproduction and which are hardly reducible to behavioral disposition that gave survival benefits to our ancestors: going to the café in order to meet someone, writing poems, collecting expressionists, and so on.

(2) Another teleological account strikes me as much more promising. Its foundations are presented in Scott Sehon's *Teleological Realism* (2005) and a couple of papers.

Agent α performed action φ in order to achieve goal *A*.

According to Sehon, action explanations are irreducibly teleological (Sehon 2013, p. 736). That means that teleological explanations cannot be reduced to causal explanations. The latter was proposed by Alfred Mele (cf. Mele 2003, pp. 55 and 58-60)[9]:

> Agent α φ'd in order to achieve *A* if and only if α's φing was caused by α's desire for *A* and her belief that φing fulfills her desire for *A*.

That is, I went to the café in order to meet my friends if and only if my desire to meet my friends and my belief that I could meet my friends if I went to the café caused it that I went to the café. Being a causal account, the causally reducible account faced the same objections as the standard view.

The irreducibly teleological account instead is freely adapting Davidson's *Principle of Charity* which says that, in order to interpret both an agent's linguistic and non-linguistic behavior, we cannot help striving for a theory of the agent's behavior that makes her as rational as possible[10]. This holds true for both the choice of goal and the choice of the means, i.e. the actions to achieve these goals. Sehon has framed two rationality principles.

> "(R_1) Agents act in ways that are appropriate for achieving their goals, given the agent's circumstances, epistemic situation, and intentional states.
>
> (R_2) Agents have goals that are of value, given the agent's circumstances, epistemic situation, and intentional states" (Sehon 2005, p. 139; cf. 2007, p. 159).

In an imperative form the first principle can be already found in Franz Brentano:

> "Choose those means that will actually bring about the end" (Brentano 1889/2009, p. 6 (§ 17)).

There the second principle firstly reads as follows:

> "Choose an end that can reasonably be thought of as being attainable" (Brentano 1889/2009, p. 7 (§ 17)).

But following the *Nicomachean Ethics*[11], this immediately turns into

---

[9] The reductive approach was initiated by an observation which Davidson made in the introduction to his *Essays on Actions and Events*: "I accept the view that teleological explanation of action differs from explanation in the natural sciences in that laws are not essentially involved in the former but hold that both sorts of explanation can, and often must, invoke causal connections" (Davidson 1980, p. xii).

[10] Davidson 1970a, p. 221-222: "[W]e cannot intelligibly attribute any propositional attitude to an agent except within the framework of a viable theory of his beliefs, desires, intentions, and decisions. [...] Global confusion, like universal mistake, is unthinkable, not because imagination boggles, but because too much confusion leaves nothing to be confused about and massive error erodes the background of true belief against which alone failure can be construed. [...] In our need to make him [i.e. an agent; G.L.] make sense, we will try for a theory that finds him consistent, a believer of truths, and a lover of the good (all by our own lights, it goes without saying)."

[11] Aristotle, *Nicomachean Ethics*, VI, 8, 1141b 12-14: "The person unqualifiedly good at deliberation is the one who tends to aim, in accordance with his calculation, at the best of the goods for a human being that are achievable in action".

"[C]hoose the best among the ends that are attainable: this is the only adequate answer" (Brentano 1889/2009, p. 7 (§ 17))[12].

Dealing with rational willing, Husserl adopts Brentano's imperative in his *Lectures on Ethics and Value Theory*:

"Do the best that is attainable!" (Husserl 1914/1988, p. 153 (§ 21); my translation)

or

"Do the best among the attainable good within your current [i.e., at the moment of choice] overall practical sphere!" (Husserl 1914, p. 142 (§ 20); my translation).

We also find this rule at the end of Davidson's *How is Weakness of the Will Possible?*, where it is called the *Principle of Continence*:

"[P]erform the action judged best on the basis of all available relevant reasons" (Davidson 1070b, p. 41).

The agent is required to perform the action which, considered in the light of all available reasons, seems her to be the best and most valuable one. Nevertheless it is questionable whether such a principle, imperative in form, is at all reasonable. It would be reasonable to require the performance of the action which has been judged to be the best only if it is possible to break the rule by performing, say, the next best action.

By contrast, Leibniz's *Principle of the Best* says that humans always strive for what appear best to them[13]. The human will cannot do other than aiming at what, all things considered, seems to be the best. It is impossible for the will to fall short of this measure. Thus the *Principle of the Best* provides the foundation for final explanations. Teleological explanations of this kind are value-based in a strong sense. The optimum supplies us with the connection between the explanatory reason and the action. It separates the reason for which the action was performed from the other reasons (see Leibniz 1714/1997, p. 187 (§§ 54-56)).

Since final explanations are value-based in a principle-of-the-best-style, the *Principle of Charity* in Davidson's version does not meet all requirements for ascertaining the explanatory reason for acting. A superlative variant of the maxim is needed: "Given two interpretations of a person's behavior, choose the one that makes her most rational!" From this Sehon obtains two principles for a teleological explanation according to which an agent ϕ'd in order to ψ:

"(I₁) Find a ψ such that ϕing is optimally appropriate for ψing, given a viable theory of the agent's intentional states and circumstances" (Sehon 2005, p. 146).

---

[12] See also Brentano 1889/2009, p. 77: "Never choose anything less than the best that is attainable".

[13] Leibniz 1685/1997, p. 39 (§ 30): "God has decreed that the will shall always tend to the apparent good, expressing or imitating the will of God in certain particular respects, in regard to which this apparent good always has something of the genuine good in it. By virtue of this decree God determines our will to the choice of that which appears the best, without necessitating it in the least."

"(I_2) Find a ψ such that ψing is the most valuable state of affairs toward which ϕing could be directed, given a viable theory of the agent's intentional states and circumstances" (Sehon 2005, p. 147; cf. 2007, pp. 163-165).

The first principle requires looking for the goal whose realization the action to be explained represents the best strategy for. The second principle requires searching the most valuable goal that can be achieved by means of the action to be explained. With both principles the action to be explained has to be considered in the light of a background theory of the agent's intentional states and circumstances; all by the interpreter's light, of course[14]. A teleological explanation will be conclusive if it sufficiently accounts for both principles.

However, it takes a *Principle of Simplicity* and a *Principle of Conservativeness* to complement those aforementioned principles. Here epistemic values such as consistency, diachronic coherence, explanatory power, and simplicity come into play. Here is an example. Suppose that when Sally pressed the cancel key of her laptop she deleted the file of her philosophy paper which was to be submitted immediately. Did Sally press the cancel key in order to delete the file? It might be that, for deleting the file, pressing the cancel key is the most appropriate means, and furthermore, owing to the circumstances, deleting the file might be the most valuable goal Sally can achieve by pressing the cancel key of her laptop. But an explanation of Sally's behavior along these lines would fail to deliver an overall theory of the agent's behavior that makes her as rational as possible, both in the actual and nearby counterfactual circumstances. Those who know Sally for quite some time may be certain that the paper on the file was important to her and she does not tend to act impulsively or rashly. Given that she thought that by pressing the key she could save the file, her behavior becomes comprehensible as being a mistake.

From the foregoing it follows that one can interpret a behavior as a goal-directed and rational action only if it does not compel us to one of the subsequent consequences: a massive revision of the background theory, a significant complication of the explanation of the agent's behavior, or a rationalization of the agent's behavior at the price of making her significantly less rational altogether. A teleologist is thus not compelled to consider a behavior goal-directed when it is not. The teleological epistemology allows for distinguishing goal-directed, rational behavior from both irrational behavior and from any behavior that is not intentional or no action at all.

Teleological explanations support counterfactuals about nearby possible worlds. This is an important point because thereby teleological explanation accomplishes what otherwise is attributed to causal explanation. I went to the café in order to meet my friends. Ceteris paribus, if meeting my friends had required going to the cinema, I would have gone to the cinema. Ceteris paribus, if I had not had the goal of meeting my friends, I would not have gone to the café. Or else: Ceteris paribus, if by going to the café I could have achieved another goal which I rated as most valuable I would have not gone to the café in order to meet my friends. To put it in a general form:

Agent α ϕ'd in order to *A*.

---

[14] For action explanations considered in the light of the agent's circumstances see also Tanney 2005.

> Ceteris paribus, if by ϕing agent α could have achieved a more valuable goal B, α would have ϕ'd in order to B.
> Ceteris paribus, if for agent α ψing would have been the optimally appropriate means for achieving goal A, α would have ψed in order to A.

Different teleological explanations of the very same behavior support different counterfactuals (cf. Sehon 2005, pp. 157-160; and 2007, pp. 167-169). The value-based teleological account provides a method for ascertaining the reason for which the agent acted and which explains her action. Consequently, it successfully takes up the Davidsonian Challenge.

5. *Objections and Replies*

The causalist standard view of action explanation considers as reason explanation those causal antecedents, i.e. desires and beliefs, causing the action in the right way. This theory seems to have a remarkable advantage. The explanatory reason being the cause of an action provides the correct explanation with an, as it were, metaphysical anchoring. The standard view draws a metaphysical distinction between the explanatory reason of an agent's action and reasons of which each would merely justify her behavior (cf. D'Oro 2007, p. 17-18; 2012, p. 213; and in contrast cf. Mele 2010). Nevertheless it has to be conceded firstly that this metaphysical anchoring does not help ascertain the correct explanatory reason for an individual action, i.e. give the correct action explanation. However, if exactly this metaphysical anchoring is at stake, the causalist is not better off than the teleologist because it is in the teleologist's option to claim that the correct teleological explanation is metaphysically anchored in the irreducible teleological fact that the behavior of the agent was directed by her toward that particular end or toward the realization of that value (Löhrer and Sehon 2016, p. 90 and pp. 92-93). I would secondly like to stress again that all attempts to give the normative expression "caused in the right way" an interpretation in purely causal terms seem to have failed up to now.

The causal theory of action explanation draws on an inference to the best explanation, in the absence of a better alternative. Teleological action explanation tries to bring forward such an alternative. An agent ϕ'd in order to A if, considered in the light of a reliable theory of the agent's intentional states and circumstances, A was the best, i.e. the most valuable end that she could achieve by ϕing and ϕing was her best means to achieve A. Nevertheless, there seem to be serious difficulties with value-based explanations.

(1) It is impossible to ascertain an optimum if the values to be compared and ranked are incomparable. But with regard to the very same observable behavior this should be out of the question. Indeed it might be that, considered in the light of a background theory of the agent's intentional states and circumstances, it is impossible to figure out

whether a visit to a concert or the repair of a washing machine is of higher value. But it is hard to imagine that both goals could be brought about by the same behavior[15].

(2) It is impossible to ascertain the explanatory reason if two alternate teleological reason explanations can refer in equal measure to an 'all-in-all' best means for an 'all-in all' best goal, where both explanations fit the background theory equally well and also meet the requirements of simplicity, conservativeness, and coherence, and where it is guaranteed that only one of the reasons explains the behavior teleologically. That might be an extremely rare case. Indeed it might be that a certain behavior is optimally appropriate to achieve two entirely different ends. Though, additional circumstances and counterfactual situations will normally still allow both cases to be distinguished. But if all such rationalizing evidence is exhausted without a verdict about the explanatory reason being reached, then the causalist faces the same problem. Since the causalist cannot read the agent's mind or observe the causal chain starting from various brain states and leading to her behavior, the causalist is forced to exercise the same procedure as the teleologist, even if she wants to understand it as a heuristic method.

(3) Here is a third objection. When introducing a value-based teleological account in addition to causal explanations, we get another type of explanation which nearly cannot apply to anything but actions (cf. Sehon 2005, ch. 13). From this point of view causalism is more parsimonious, and from an economic vantage point causalist theories should even be preferred. However, if the objection I raised against causalism in the previous paragraph is convincing, this also pertains to the reducible teleological account. Thus we would have to accept that explanations of events which are actions and of events which are not are mutually irreducible. They explain different things in different ways[16].

References

- ARISTOTLE, *Nicomachean Ethics*, ed. and transl. R. Crisp, Cambridge University Press, Cambridge UK 2000.
- F. BRENTANO (1889/2009), *The Origin of the Knowledge of Right and Wrong*, ed. R.M. Chisholm, transl. R.M. Chisholm and E.H. Schneewind, Routledge, London 2009.
- R. CHANG (1997), *Introduction*, in EAD. (ed.), *Incommensurability, Incomparability, and Practical Reason*, Harvard University Press Cambridge MA 1997, pp. 1-34, 254-261.
- W. CHILD (1994), *Causality, Interpretation and the Mind*, Clarendon Press, Oxford UK 1994.
- M. CAVELL (2004), *Introduction*, in D. DAVIDSON, *Problems of Rationality*, Clarendon Press, Oxford UK 2004, pp. xiii-xx.
- D. DAVIDSON (1963), *Actions, Reasons, and Causes*, in ID., *Essays on Actions and Events*, Clarendon Press, Oxford UK 1980, pp. 3-19.

---

[15] See also Ruth Chang's objections to arguments for the incomparability of values (Chang 1997).

[16] For proof reading of my English I wish to thank Beate Hampe.

- D. DAVIDSON (1970a), *Mental Events*, in ID., *Essays on Actions and Events*, Clarendon Press, Oxford UK 1980, pp. 207-227.
- D. DAVIDSON (1970b), *How is Weakness of the Will Possible?*, in ID., *Essays on Actions and Events*, Clarendon Press, Oxford UK 1980, pp. 21-42.
- D. DAVIDSON (1973), *Freedom to Act*, in ID., *Essays on Actions and Events*, Clarendon Press, Oxford UK 1980, pp. 63-81.
- D. DAVIDSON (1974), *Psychology as Philosophy*, in ID., *Essays on Actions and Events*, Clarendon Press, Oxford UK 1980, pp. 229-239.
- D. DAVIDSON (1976), *Hempel on Explaining Action*, in ID., *Essays on Actions and Events*, Clarendon Press, Oxford UK 1980, pp. 261-275.
- D. DAVIDSON (1978), *Intending*, in ID., *Essays on Actions and Events*, Clarendon Press, Oxford UK 1980, pp. 83-102.
- D. DAVIDSON (1978), *Problems in the Explanation of Action*, in ID., *Problems of Rationality*, Clarendon Press, Oxford UK 2004, pp. 101-116.
- D. DAVIDSON (1984), *Expressing Evaluations*, in ID., *Problems of Rationality*, Clarendon Press, Oxford UK 2004, pp. 19-37.
- G. D'ORO (2007), *Two Dogmas of Contemporary Philosophy of Action*, in "Journal of the Philosophy of History", 1 (1/2007), pp. 10-24.
- G. D'ORO (2012), *Reasons and Causes. The Philosophical Battle and the Meta-philosophical War*, in "Australasian Journal of Philosophy", 90 (2/2007), pp. 207-212.
- A. GOLDMAN (1970), *A Theory of Human Action*, Prentice-Hall, Englewood Cliffs NJ 1970.
- J. HORNSBY (1980), *Actions*, Routledge & Kegan Paul, London 1980.
- E. HUSSERL (1914/1988), *Vorlesungen über Ethik und Wertlehre – 1908-1914* (Husserliana XXVIII), ed. U. Melle, Kluwer, Dordrecht-Boston-London 1988.
- G. KEIL (2007), *What Do Deviant Causal Chains Deviate From?*, in C. LUMER-S. NANNINI (ed.), *Intentionality, Deliberation and Autonomy. The Action-Theoretic Basis of Practical Philosophy*, Ashgate, Aldershot 2007, pp. 69-90.
- G.W. LEIBNIZ (1685/1997), *Discourse on Metaphysics*, in ID., *Philosophical Writings*, ed. G.H.R. Parkinson, Everyman, London 1997, pp. 18-47.
- G.W. LEIBNIZ (1714/1997), *Monadology*, in ID., *Philosophical Writings*, ed. G.H.R. Parkinson, Everyman, London 1997, pp. 179-194.
- G. LÖHRER (2006), *Abweichende Kausalketten, abwegige Handlungsverläufe und die Rückkehr teleologischer Handlungserklärungen*, in "Deutsche Zeitschrift für Philosophie", 54 (5/2006), pp. 785-800.
- G. LÖHRER (2008), *Alltagspsychologische Handlungserklärungen, Kausalität und Normativität*, in "Internationale Zeitschrift für Philosophie", 17 (1/2008), pp. 76-100.
- G. LÖHRER-S. SEHON (2016), *The Davidsonian Challenge to the Non-Causalist*, in "American Philosophical Quarterly", 53 (1/2016), pp. 85-95.
- A.R. MELE (2003), *Motivation and Agency*, Oxford University Press, New York NY 2003.
- A.R. MELE (2010), *Teleological Explanations of Actions: Anticausalism vs. Causalism*, in J. AGUILAR-A. BUCKAREFF (ed.), *Causing Human Actions: New Perspectives on the Causal Theory of Action*, MIT Press, Cambridge MA 2010, pp. 183-198.
- A.R. MELE (2013), *Actions, Explanations, and Causes*, in G. D'ORO (ed.), *Reasons and Causes: Causalism and Non-causalism in the Philosophy of Action*, Macmillan, Basingstoke 2013.

- R.G. MILLIKAN (1993), *Explanation in Biopsychology*, in EAD., *White Queen Psychology and Other Essays for Alice*, MIT Press, Cambridge 1993, pp. 171-192.
- G.F. SCHUELER (2001), *Action Explanations: Causes and Purposes*, in B.F. MALLE-L.J. MOSES-D.A. BALDWIN (ed.), *Intentions and Intentionality. Foundations of Cognition*, MIT Press, Cambridge MA 2001, pp. 251-264.
- J. SEARLE (1983), *Intentionality. An Essay in the Philosophy of Mind*, Cambridge University Press, Cambridge UK 1983.
- S. SEHON (1997), *Deviant Causal Chains and the Irreducibility of Teleological Explanation*, in "Pacific Philosophical Quarterly", 78 (2/1997), pp. 195-213.
- S. SEHON (2005), *Teleological Realism. Mind, Agency, and Explanation*, MIT Press, Cambridge MA 2005.
- S. SEHON (2007), *Goal-Directed Action and Teleological Explanation*, in J.K. CAMPBELL-M. O'ROURKE-H.S. SILVERSTEIN (ed.), *Topics in Contemporary Philosophy: Causation and Explanation*, MIT Press, Cambridge MA 2007, pp. 155-170.
- S. SEHON (2013), *Teleology*, in H. PASHLER (ed.), *Encyclopedia of the Mind*, Sage Publications, Beverly Hills CA 2013, pp. 735-737.
- J. TANNEY (2005), *Reasons-Explanations and the Contents of the Mind*, in "Ratio" (new series), 18 (2005), pp. 338-351.
- G.M. WILSON (1989), *The Intentionality of Human Action*, revised and enlarged Edition, Stanford University Press, Stanford 1989.