

Davidson on Justification and Rationalization

LILIAN O'BRIEN*

Abstract: It is widely assumed that action explanation renders intentional actions rationally intelligible in light of an agent's beliefs and desires. Donald Davidson shares this assumption. However, I argue that his theory of action explanation fails to accommodate this feature of action explanation. Moreover, this failure offers strong reasons to abandon causalism about action explanation.

Key words: justification, rationalization, intentional action, reason for action

Resumen: Suele suponerse que la explicación de las acciones hace a las acciones intencionales inteligibles racionalmente a la luz de las creencias y deseos del agente. Donald Davidson comparte este supuesto. Sin embargo, defenderé este rasgo de la explicación de la acción no encuentra acomodo en su teoría de la explicación de las acciones. Más aún, esta insuficiencia proporciona razones poderosas para abandonar el causalismo sobre la explicación de las acciones.

Palabras clave: justificación, racionalización, acción intencional, razón para la acción.

Introduction¹

Donald Davidson, among others, plausibly assumes that rationalization² renders an intentional action rationally intelligible in light of the reasons on which the agent acted.³ With this assumption in mind, it seems that one task of a theory of rationalization is to provide an account of how rationalization makes intentional actions rationally intelligible. In the following, I will argue that although Davidson acknowledges the central role of rational intelligibility to rationalization, his view⁴ does not provide an analysis of rationalization that accounts for this feature.

Fecha de recepción: 19 mayo 2005. Fecha de aceptación: 28 octubre 2005.

* Lilian O'Brien, Ph.D., Visiting Assistant Professor, Department of Philosophy, Vassar College, Poughkeepsie, NY 12604, USA; liobrien@vassar.edu.

1 I would like to thank the members of a reading group at Brown University who read an earlier version of this paper, Aaron Boyden, Christopher Kane, Jaegwon Kim, Alyssa Ney, Rebekah Rice and Nick Treanor. Thanks also to Antti Kauppinen for comments on an earlier draft, and to members of the audience at SEFA 2004 for helpful comments and questions.

2 This is taken to be the explanation of intentional actions by reasons.

3 In «Actions, Reasons and Causes» (1963) Davidson says: «A reason rationalizes an action only if it leads us to see something the agent saw, or thought he saw, in his action – some feature, consequence, or aspect of the action the agent wanted, desired, prized, held dear, thought dutiful, beneficial, obligatory, or agreeable.» In «Intending» (1978) he says, «To serve as reasons for an action, beliefs and desires need not be reasonable, but a normative element nevertheless enters, since the action must be reasonable in the light of the beliefs and desires...»

4 Clearly, Davidson's views changed over time, but I will here concentrate on his 1963 view. Although there is no exegetical treatment here of such later work as his 1978 paper «Intending» nor of his 1987 paper «Problems in the Explanation of Action», the argument of this paper applies equally forcefully against his views expressed in these later papers. Demonstrating this goes well beyond what I can discuss here.

In Section 1, I will discuss the widely accepted view that rationalization tracks⁵ «motivating reasons». These are psychological states of the agent, especially beliefs and desires. I argue that we face a dilemma: either rationalization tracks motivating reasons or it tracks normative reasons. If the former, then rationalization does not necessarily make intentional actions rationally intelligible, if the latter, then it is obscure how rationalization works in cases where reasons that lack justificatory force in fact are nevertheless explanatory.

In Section 2, I motivate what, I believe, is the most plausible path out of the dilemma. My claims in this section do not amount to a theory of rationalization, nor do I offer a solution to the problem raised in Section 1. These remarks are intended as a broad outline of the kind of solution that must be offered to the problems discussed earlier in the paper.

Section 1: The Appeal of Motivating Reasons

It is widely assumed that rationalization renders an intentional action rationally intelligible in light of the reason that justified the action. This assumption simply reflects how rationalization appears to work in everyday contexts. By examining such contexts, we can show that, typically, we do not take ourselves to have successfully completed the rationalizing project until we come to see that the agent was, in some sense, rationally justified in undertaking what she did undertake to do. Suppose that I see you open the window and am told that the reason why you are doing that is because you want to cool the room. Now, suppose that I believe that it is hotter outside than inside, and do not believe of the window-opening that it is a good means to the end of cooling the room. In such circumstances, I may be puzzled as to why you would open a window to achieve your end. My puzzlement indicates that the expectation in rationalization is that we come to see the agent's action as rationally intelligible in means-end terms. In fact, it is highly usual in such circumstances to pursue the rationalizing enquiry by asking why it is that you are opening the window, given that it is hotter outside than inside. Let's suppose that I discover that you didn't know that it is hotter outside than inside. With this information about your beliefs, it seems that I can see why, given what you took to be the case, you thought that opening the window was a good means to cooling the room.

An important aim of any theory of explanation is to isolate what it is about the relationship between the *explanans* phenomenon and the *explanandum* phenomenon that allows the former to explain the latter.^{6 7} For example, we might ask what it is about the relationship between a baseball flying through the air at a certain speed and the breaking of a window that results in the fact that the former event can explain the latter. The answer seems to be that the former caused the latter, and it is in virtue of its causing it that it can explain it.⁸ From the point of view of analyzing rationalization,

5 I will sometimes talk about explanation as «tracking» things. I follow Kim in this usage. I assume that explanation has an epistemic and a metaphysical component, where the former consists of statements that describe the latter in a special explanatory kind of way. On these minimal and widely accepted assumptions, I am going to help myself to the notion of «tracking» to capture the dependence of the epistemic component of explanation on the metaphysical component. For discussion of these and related issues, see Kim 1994, 1999.

6 The *explanans* phenomenon is that which does the explaining, the *explanandum* phenomenon is that phenomenon that is being explained.

7 I will maintain a realist view of explanation throughout. Roughly, explanatory realism is the view that there are features independent of the descriptive component of the explanation that are «mapped» or tracked by the description. Explanation is not just a matter of «internal» relations among propositions. See Kim, 1987.

8 Of course, much more needs to be said to show why causation is explanatory.

we must understand what it is about reasons that allows them to rationalize actions. From the foregoing brief discussion of how rationalization works in everyday contexts, it is clear that the explanatory relation holding between the reason and the action must have some sort of justificatory component. The rationalizer must come to see the agent's action as, in some sense, justified by her reason for action. It would clearly be very problematic to construe rationalization just as justification, or in other words, to claim that the reason explains precisely because it justifies the action. This is because there are reasons that do not objectively justify actions, but nevertheless rationalize them. Opening the window to cool the room in spite of the fact that it is hotter outside than inside is a case in point. Consequently, we must accommodate both the justificatory aspects of the relation that holds between the reason and the action and the fact that «bad reasons», reasons that do not objectively justify actions, may nevertheless rationalize them perfectly well.

A distinction between «normative reasons» and «motivating reasons» is often thought to help with this problem.⁹ Normative reasons are taken to have objective normative force and they are taken to justify actions. Motivating reasons motivate the agent to act in one way or another, and they may or may not have objective normative force. They nevertheless «motivate» the agent to act, and we can talk about the agent's reasons without committing ourselves to any claim about whether or not the action is objectively justified. However, motivating reasons are not devoid of normative content, and are, presumably, rationally intelligible.

Motivating reasons have proven very difficult to understand, so it is not at all clear what the foregoing shift amounts to. For our purposes, we will confine ourselves to an examination of one view of rationalization that relies on motivating reasons, and we will eschew the much more difficult task of trying to understand motivating reasons as they are presented throughout the literature. This view to be examined is that of Donald Davidson in «Actions, Reasons and Causes». In this paper, he takes a motivating reason to be at least the combination of a pro-attitude and a belief.¹⁰ Davidson claims that an intentional action is explained, not just by any combination of pro-attitude and belief, but is explained by a primary reason composed of a pro-attitude and belief that is constrained in the following way:

C1: R is a primary reason why an agent performed the action A under the description d only if R consists of a pro attitude of the agent towards actions with a certain property, and a belief of the agent that A, under the description d, has that property.¹¹

I take Davidson's position in «Actions, Reasons and Causes» to be that a reason that conforms to C1 and is causally efficacious in producing an intentional action is a reason that can rationalize an intentional action. The added requirement that a reason that conforms to C1 must also be causally efficacious in the production of the action arises from what has become known as Davidson's challenge. In «Actions, Reasons and Causes» Davidson challenges the non-causalist to be able to distinguish between the reasons for action that the agent may have had but for which the agent did not act, and the reason or reasons for which the agent ultimately acted. Davidson's view is that it is the causal efficacy in the production of action that distinguishes *the reason or reasons for which* the

9 See Dancy, 2000, for a penetrating discussion of the distinction.

10 Davidson, 1963; James Dreier says «Motivating reasons are psychologically real ... A motivating reason that you have is an empirical property that you bear ...» Dreier, 2001. See also Michael Smith, 1987, 1998.

11 Davidson, 1963, p. 5.

agent acted from *the reason or reasons for* action that the agent had but didn't act on. Hence, he argues that C1 must be bolstered by a further condition, C2.

C2: A primary reason for an action is its cause.¹²

At this point, two questions arise. First, if we adopt the view that rationalizations are tracking motivating reasons, what feature or features of those reasons are doing the explanatory work in rationalization? Second, given our assumption that rationalization renders intentional actions rationally intelligible, can we capture this fact about rationalizations with our analysis of rationalization in terms of motivating reasons construed as Davidson construes them? Davidson has a ready answer to the first question, namely, the reason explains the action because it causes it.¹³ As for the second question, I will argue that Davidson fails to preserve this feature of rationalization.¹⁴

Let us consider the following rather bizarre case of Sally. Suppose that Sally wants to alleviate her thirst and has the belief that opening the window is a means to alleviating her thirst. We can then construct a primary reason using this desire and belief that conforms to C1. In addition, we can stipulate that it is causally efficacious. According to Davidson's 1963 view, this reason rationalizes Sally's action of opening the window because it satisfies the necessary and sufficient conditions (C1 and C2) on being the «reason for which» the agent acts. However, when we put ourselves in the position of a rationalizer who wants to know why Sally is opening the window, being told that she is alleviating her thirst or attempting to alleviate her thirst in this action does not make her action rationally intelligible. Indeed, we can imagine ourselves being very bewildered in such a case. This suggests that Davidson's conditions on intentional action and the reason for which are not sufficient for specifying the conditions in which a reason for which an agent acts is sufficient to rationalize that action. Specifically, it seems clear that the assumption that rationalization makes actions rationally intelligible is violated if we take C1 and C2 to be both necessary and sufficient for a reason that rationalizes an action.

Let's call the case of Sally's reason, which conforms to Davidson's C1 and C2, and thus should, by Davidson's lights, be sufficient to rationalize her action of window-opening, the «crazy reason» problem (as opposed to the «bad reasons» problem alluded to above, wherein we must accommodate reasons that rationalize although they don't objectively justify actions). The idea behind the «crazy reason» problem is just that motivating reasons are taken to be empirically real psychological states of the agent, which don't obviously have features that preclude their arrangement in a rationally incoherent fashion.¹⁵ If it is possible to have a primary reason such as Sally's, then a tension arises between assuming that rationalization makes actions rationally intelligible and holding that motivating reasons rationalize intentional actions. It seems that we must either find a constraint on motivating reasons or we must give up our assumption that rationalization makes actions rationally intelligible. As the assumption is not just something that

12 Davidson, 1963, p. 12.

13 Dreier and others point out that tracking motivating reasons does not entail tracking causal relations, but it is also a little obscure what else we could be tracking. See Dreier (1997).

14 This claim is not novel, of course, but my argument for it is different from those in the literature that I am aware of.

15 I take the rational constraints on belief-desire pairs that arise in Davidson's interpretationism to be the result of epistemic constraints on the interpreter and the interpreted that do not derive from the metaphysical nature of the actual psychological states of the agent in question. I am here taking Davidson's interpretationism and his action theory separately.

Davidson appears to subscribe to, but is reflected in everyday cases of rationalization¹⁶, it is not the obvious candidate for abandonment.

The Problem with the Move to Motivating Reasons

The introduction of «motivating reasons» was meant to capture the power of reasons in rationalization to render actions rationally intelligible while factoring in the detachment from the objective normative force of those reasons. On Davidson's version of this outlook, the feature of rationalization that is that it makes actions rationally intelligible is not preserved. Indeed, it seems that we face a dilemma: either reasons that objectively justify actions rationalize them, in which case we are without an account of reasons that rationalize although they lack justificatory force (the «bad reasons» problem), or motivating reasons rationalize, in which case rationalization is no longer constrained by a rational intelligibility requirement (the «crazy reasons» problem).

Section 2: The Relegation of Causation

How then are we to resolve the dilemma that we face? In the following section, I will urge that we tread a particular path to our solution, but as aforementioned, what I say will not amount to either a full solution to the problem raised.

The notion of motivating reasons is a disputed one, and I have not, so far, devoted time to exploring the possibilities for fleshing it out. My main aim has been to criticize the Davidsonian view, which fails to show how we can constrain motivating reasons to secure a full analysis of rationalization. A pertinent consideration is that if we were, in an everyday context, to be told that Sally was opening the window to alleviate her thirst, we would be likely to embark upon an exploration of Sally's background beliefs to try to understand why she thinks that window-opening is a good means to her end. For example, if we were told that Sally believes that in opening the window she sends a signal to someone to send her a drink of water, we may come to see why Sally thought that she was rationally justified in doing what she did. Of course, we may not believe of many of Sally's beliefs that they are true, but it would be possible for us to reconstruct her reasoning, or, if you like, the rational outlook that she had on her action. We could thereby make her action rationally intelligible to us.

In light of this fact, we might amend Davidson's C1 to reflect this holistic character of rationalization by, for example, stipulating that in cases where rational intelligibility does not result from coming to know what the causally efficacious primary reason was, we can proceed to enquire into background beliefs and desires to complete the process of rationalization. What does this tell us about the nature of rationalization? First, it reinforces the point already made about Sally's «crazy reason» case, namely, that we must see the action as, in some sense, justified, if we are to rationalize it. Second, it shows that whatever the metaphysical character of motivating reasons, if they are to qualify for the role of *explanans* in rationalization, they must be able to play a justificatory role, not only as adequate, albeit subjective, justifications for actions, but as part of a potentially large network of justificatory relations that feed the rational support that an agent may have for any given action. In light of this, the poverty of the purely causalist approach to motivating

¹⁶ For example, in the crazy reason case, we would typically be dissatisfied with the «answer». In fact, we would typically refuse to let the matter rest and begin enquiring into the agent's other beliefs to try to discern why they did what they did.

reasons in Davidson's analysis seems clear, and strongly suggests that we must accommodate the «bad reasons» problem without appeal to causation. This is not yet a refutation of causalism with respect to rationalization, as there may still be some explanatory role left for causation. In the following, I will urge a skepticism about such a possibility. I will briefly discuss belief rationalization as a potential model on which to understand action rationalization. It recommends itself as a model because it has many commonalities with the rationalization of action, and it appears to accommodate «bad reasons» without appeal to causation.

When we wish to rationalize an agent's belief, we may enquire into her justification for her belief in an effort to understand it. As with the rationalization of action, it may require an examination of a large number of the agent's background beliefs and the justificatory relations that she takes to hold among them. In both cases the aim is to understand the rational intelligibility of the agent, and in neither case is the rationalizer committed to the objectivity of the agent's justification. Both processes accommodate «bad reasons». True, in one case we rationalize an agent's belief and in the other we rationalize her action. However, what seems crucial to this latter project is that we understand the rationality of the agent's action in light of her judgment that the action is «the thing to do». Consequently, in both cases it seems that we are assessing the rationality of a judgment of the agent's, and thus, the difference does not seem to be of great import.

It is interesting that, in spite of such a strong similarity between the two types of rationalization, it is not suggested of the rationalization of beliefs that it is a type of causal explanation, nor is it suggested that an appeal to causation would provide a solution to the problem of understanding rationalization of beliefs in the case of «bad reasons» for belief. And yet, Davidson's challenge seems to be equally applicable here. After all, what is the difference between the *reasons for belief* that the agent had, and the *reasons for which* she ultimately formed the belief? If there is no philosophically significant difference between the types of rationalization, why has Davidson's challenge not been applied to the process of belief formation? I think that the reason for this is that everyday rationalization is a practice in which the rationalizer simply assumes that the agent fulfills the causal conditions on being an epistemic agent. If this is right, it seems that rationalization of belief is a process that presupposes certain facts about the causal order, specifically those facts concerning whether or not the rationalizer's deliberation is causally efficacious in the production of beliefs. If it is a presupposition of the enquiry that the causal order is a certain way, then the enquiry cannot be concerned with tracking and describing that part of the causal order.¹⁷ If this picture of belief-rationalization is roughly correct, it seems to combine a focus on the justificatory features of reasons while accommodating «bad reasons», and it seems to do this without appeal to causation.

The foregoing is a rough and speculative picture of how belief-rationalization works, and much more needs to be said to flesh it out. However, what it suggests is that, given the similarities between the rationalization of actions and beliefs, and given, as I hope to have established in the earlier part of the paper, the prominence of the justificatory aspect of «motivating reasons» in the rationalization of action, we should attempt to resolve our dilemma, not by giving up on rational intelligibility and embracing causalism, nor by ignoring the fact that causalism does not explain rational intelligibility, but by looking at the rationalization of belief as a potential source of insight into the nature of the rationalization of action. Belief-rationalization seems to prioritize the tracking

17 I discuss this elsewhere in much greater detail in «A Non-Causalist Theory of Rationalization», Brown University Ph.D. dissertation, May 2006.

of justificatory relations, while accommodating «bad reasons» and it also seems to relegate causation to the presuppositions of the rationalizing enquiry. Even if we are at a loss to fully understand the process of rationalization in the case of beliefs, it does present itself as a starting point for a resolution to the problem of accommodating justification in rationalization.

Conclusión

I have not attempted to offer either an analysis of rationalization, or a solution to the problem posed by the dilemma above. My aims have been to highlight the prominent place of justification in rationalization, and to establish the poverty of Davidson's approach in light of this. What I am proposing is that we take the case of belief rationalization as an interesting place to start in trying to understand the nature of the rationalization of actions. It suggests itself because it appears to accommodate the prominent place of justificatory relations, the problem of «bad reasons» and it seems to do so without relying on causation.

Bibliography

- Dancy, Jonathan. (2000) *Practical Reality*. Oxford: Oxford University Press.
 «Precis of *Practical Reality* and Replies», Book Symposium, *Philosophy and Phenomenological Research*, Vol. LXVII, No. 2.
- Davidson, Donald. (1963) «Actions, Reasons and Causes», reprinted in *Essays on Actions and Events*, Oxford and New York: Oxford University Press, 1980.
- (1978) «Intending», reprinted in *Essays on Actions and Events*.
 - (1980) *Essays on Actions and Events*. Oxford and New York: Oxford University Press.
 - (1987) «Problems in the Explanation of Action», reprinted in *Problems of Rationality*, Oxford and New York: Oxford University Press, 2004.
- Dreier, James. (1997) «Humean Doubts about Categorical Imperatives», in *Ethics and Practical Reasoning*, Garrett Cullity and Berys Gaut, eds., Oxford: Clarendon Press. Reprinted in *Varieties of Practical Reasoning*, (2001) (ed.) Elijah Millgram.
- Kim, Jaegwon. (1987) «Explanatory Realism, Causal Realism, and Explanatory Exclusion», *Midwest Studies in Philosophy*, Vol. 12 (1987): 225-240. Reprinted in *Explanation*, ed. David-Hillel Ruben, Oxford: Oxford University Press, 1993.
- (1994) «Explanatory Knowledge and Metaphysical Dependence», *Philosophical Issues: Truth and Rationality*, pp. 51-69.
 - (1999), «Hempel, Explanation, Metaphysics», *Philosophical Studies* 94, pp. 1-20.
- Smith, Michael. (1987) «The Humean Theory of Motivation», *Mind*, Vol. 96, No. 381, pp. 36-61.
- (1998) «The Possibility of Philosophy of Action», *Human Action, Deliberation and Causation*, eds. J. Bransen and S.E. Cuypers, pp. 17-41.

