

Reasons for Action

Pamela Hieronymi
hieronymi@ucla.edu

Abstract: Donald Davidson opens ‘Actions, Reasons, and Causes’ by asking, ‘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?’ His answer has generated some confusion about reasons for action and made for some difficulty in understanding the place for the agent’s own reasons for acting, in the explanation of an action. I offer here a different account of the explanation of action, one that, though minimal and formal, preserves the proper role for the agent’s own reasons for acting.

Donald Davidson opens his seminal ‘Actions, Reasons, and Causes’ (1980c) with the following question: ‘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?’ His answer seems to be that a reason that explains the action in this way (a reason that ‘rationalizes’ the action) also causes the action. Many have been convinced by his argument for this answer, and have thus been led to think of reasons for action as psychological states that cause action—as beliefs and desires, in particular.¹

However, the thought that reasons for action are psychological states that cause action is, at least *prima facie*, in tension with another natural thought: A great many people thinking and writing about reasons take reasons for action to be, not psychological states that cause action, but rather considerations that count in favor of acting: facts about the importance of the election, the cost of the theater ticket, or the distance to the subway. This thought seems natural, since, as Davidson himself puts it, in giving ‘the agent’s reason for doing what he did,’ we point to ‘some feature, consequence, or aspect of the action the agent wanted, desired, held dear, thought dutiful, beneficial, obligatory, or agreeable’ (1980c: 3). One might point to the fact that the action would promote the challenger or discharge a debt. It is natural to think that these features, consequences, or aspects of the action were the agent’s

¹ For Davidson, causes are always events, so, more strictly, Davidson claims that an event that is the onset of certain psychological states stands in a causal relation to an event that is an action. See, e.g., Donald Davidson, ‘Actions, Reasons, and Causes’, *Essays on Actions and Events* (New York: Oxford University Press, 1980c), 3–19 at 12–13.

reasons for acting. But these are not psychological states that cause action. They are rather facts or considerations the agent took to count in favor of acting.

There is some confusion, then, about reasons for action. The idea of a ‘reason for action’ seems to some to pick out considerations that count in favor of acting, while to others it picks out psychological states that explain action. In what follows, I first do a bit of taxonomy, attempting to tidy the field and avoid verbal dispute by stipulating some technical terms. We will see that the disputants are thinking of ‘reasons for action’ in different, but not necessarily incompatible, ways.

I then consider some approaches to explaining action, including Davidson’s own. We will see each of them fail, though for different reasons, and so we will develop criteria for an account of the explanation of action.

Finally, I will propose an embarrassingly simple, alternative account of the form of action explanation: if we care to explain an event in such a way as to make clear that it was an action done for certain reasons, we can do so by noting that the agent, for those reasons, settled for him or herself the question of whether so to act, therein intended so to act, and executed that intention in an action that was the event in question.

In its simplicity, this formal account leaves unexplained much of what others hoped to explain. Most glaringly, the account leaves unanalyzed, at its heart, the activity of the agent, in settling a question. It also insists that an intention is ‘executed’ in action, without further explicating the relation between the intention and the action. One should hope for elaboration. In fact, it may seem I have provided only a kind of *analysis* of action done for reasons, not an *explanation* of an event that is an action—perhaps my proposed form will become genuinely explanatory only once filled in.² Even so, I believe introducing this simple, formal account may be both helpful and important; in particular, it can help to clarify the

² I am especially grateful to Simon Rippon for pressing this point.

roles played by reasons, in explaining action done for reasons. Thus, I believe it can provide guidance to those who pursue the more ambitious projects.

I will close by noting that the proposed account avoids the difficulties encountered by others by adopting a different underlying account of what a reason is. Whereas others understand reasons as facts or considerations standing in some relation to events, states of affairs, actions, or attitudes (e.g., causing, explaining, or counting in favor of them), I understand reasons as relating, first, to questions (and to events, etc., *via* questions). By so understanding reasons, we can avoid the objections that led Davidson to conclude that action explanation must be a species of causal explanation, without insisting that action explanation is a species of causal explanation. The view I propose thus occupies a middle ground between Davidson and his opponents.

1. A TAXONOMY OF REASONS FOR ACTION: THREE CONTENDERS

As noted, some philosophers take reasons to be, fundamentally, *considerations that count in favor of an action or attitude* (such as the fact that it is getting late, or that no one has yet eaten dinner) (Dancy 2000; Parfit 2001; Raz 1975, 1999; Scanlon 1998).³ But taking this as our fundamental account generates difficulty.⁴

The difficulty first appears upon considering happenings or states of affairs that are not themselves activities done for reasons, such as an engine's failure. We think there are reasons that explain such happenings: the engine failed because of the extreme heat or the faulty

³ According to Allan Gibbard, to call R a reason for action is to express 'acceptance of norms that say to treat R as weighing in favor of doing' that action. See Alan Gibbard, *Wise Choices, Apt Feelings* (Cambridge, MA: Harvard University Press, 1990). According to John Skorupski, a reason for action is relation holding between a fact, a person, an action, and a time. See, e.g., John Skorupski, 'Reasons and Reason', in Garrett Cullity and Berys Gaut (eds.), *Ethics and Practical Reason* (Oxford: Oxford University Press, 1997), 345–67.

⁴ I notice another difficulty in Pamela Hieronymi, 'The Wrong Kind of Reason', *The Journal of Philosophy*, 102/9 (September 2005), 1–21. (The present paper was once half of that article; it provides further support for my preferred view, that a reason—of any sort—is a consideration that bears or is taken to bear on a question.)

construction. But the reasons that explain such events do not count in favor of them. The faulty construction does not count in favor of the engine's failure: it does not show anything good, right, desirable, appropriate, or important about the failure. On this, everyone agrees.⁵

In response, some philosophers simply divide reasons at the root, so to speak, between, on the one hand, what are called 'normative' or 'justifying' reasons—reasons which show a given action, attitude, activity, or outcome good, right, appropriate, or called for—and, on the other, what are called 'explanatory' reasons—the reasons why things happen, or why things are the way they are.⁶ These are, it might be thought, simply different senses of 'reason', or perhaps different 'kinds' of reasons, which appear in different contexts (of explanation, on the one hand, and deliberation, on the other). Only the so-called 'normative' or 'justifying' reasons count in favor of actions or attitudes; 'explanatory' reasons are simply something else.

However, having drawn this divide, complication arises in what appears to be a hybrid case: we care to explain those events that are (also) actions done for reasons.⁷ We care to explain why the engine failed, but we also care to explain why Warren spoke up and why Erin left the room. That is, we care to explain—to give 'explanatory' reasons for—actions done for reasons—done for the so-called 'normative' reasons. How are we to relate these reasons, in such an explanation?

⁵ It might show the failure likely or intelligible, but that is a different matter.

⁶ Scanlon does so in T. M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998). Some offer an account of the connection between the explanatory and 'normative.' Bernard Williams insisted that any reason for action must a potential explanation of the action. See Bernard Williams, 'Internal and External Reasons', *Moral Luck* (1981), 101-13. John Broome identifies 'normative' reasons (specifically, 'perfect' reasons) as facts that explain ought claims. See John Broome, 'Reasons', in R. Jay Wallace et al. (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (Oxford: Clarendon Press, 2004), 28–55. According to Mark Schroeder, a consideration is a 'normative' reason for action, roughly, if it is part of an explanation of why a given action would satisfy some desire. See Mark Schroeder, *Slaves of the Passions* (Oxford: Oxford University Press, 2008) 224.

⁷ The same difficulty appears when explaining attitudes held for reasons.

At first, one might think the difficulty is not so grave: perhaps, to explain an action done for a reason, we can simply appeal to the reason for which the action was done. Suppose Erin left the room because the meeting was over. One might think the meeting's end explains Erin's departure. On such a simple, first-pass account, the relation between the 'explanatory' and the so-called 'normative' reasons seems to be simple identity.

The simple account runs into trouble, once we notice the unfortunate possibility of error: perhaps Erin left the room because *she thought* the meeting was over, but it was not.⁸ Such error generates two types of difficulty.

First, if the meeting was not over, then it seems that nothing actually counted in favor of leaving—there was, in fact, no reason to leave. And yet, Erin acted for a reason (she did not act on a whim, for no reason). Thus, it seems, she acted for a reason that was no reason.⁹ To accommodate this awkward possibility, we need to be able to refer to the considerations that someone took to count in favor of an action, whether or not they actually count in favor of it—those considerations someone *treated as* 'normative' reasons. T. M. Scanlon (1998: 19) calls the considerations someone took to count in favor of acting, on which he or she acted, 'operative' reasons; Jonathan Dancy (2000), Allan Gibbard (1990: 162), Derek Parfit (2001),

⁸ Thanks are due to Jonathan Dancy, for bringing this problem into relief. See Jonathan Dancy, *Practical Reality* (Oxford: Oxford University Press, 2000) 187. I here focus on mistakes of fact. One might instead mistake what the facts count in favor of doing. Such cases generate further complication, but, I believe, can be handled in the same way I will propose handling mistakes of fact.

⁹As Dancy puts it, 'there was no reason to do what [she] did, even though [she] did it for a reason' *Ibid.* at 3. Notice the pressure both to allow as (real) reasons what others take to be reasons and to insist that the only (real) reasons are the good reasons.

and Mark Schroeder (2008) call these ‘motivating’ reasons.¹⁰ Davidson has something like this in mind when referring to ‘the agent’s reason for doing what he did.’

Second, and perhaps more troublingly, it seems that these operative reasons cannot explain the action. If Erin’s operative reason for leaving was that the meeting was over, but the meeting was not over, then we cannot appeal to the meeting’s end to explain her departure—because the meeting did not end. Something that is not the case cannot explain something that is.¹¹ To provide an explanation of Erin’s departure, one must cite some fact. The natural fact to employ is the (psychological) fact that Erin *thought* the meeting was over. Michael Smith (1994: 131) calls such psychological facts ‘motivating reasons.’¹² I will henceforth adopt Smith’s (less usual) terminology and use ‘motivating reasons’ to refer to psychological facts that explain action.

It is important to note the distance between *Smith’s* ‘motivating reasons,’ on the one hand, and Scanlon’s ‘operative reasons’ (many other philosopher’s ‘motivating reasons’) on the other. It is important, that is, to note the contrast between the reasons that seem to explain Erin’s leaving and Erin’s own reasons for leaving. The reasons that explain her action seem to be facts about her psychology. But Erin’s own reason for leaving—that which she took to count in favor of acting—had nothing to do with her psychology. She did not take

¹⁰ Scanlon’s use here differs somewhat, it seems to me, from Raz’s use of the same label in Joseph Raz, *Practical Reason and Norms* (London: Hutchinson, 1975) at 33. Schroeder provides a helpful taxonomy, in which he distinguishes, usefully, between what he calls ‘subjective normative’ reasons and ‘motivating’ reasons. Though this is an important distinction, I suppress it, here. I believe his motivating reasons are, roughly, Scanlon’s operative reasons. See Schroeder, *Slaves of the Passions* at 12–15.

Yet another category: considerations that another *ought* to have taken to count in favor of this or that—the reasons that we notice when we imaginatively enter her point of view, regardless of whether she noticed them or took them into account, and regardless of whether they are, in fact, good reasons. These seem to me better contenders for title ‘justificatory’ reasons, insofar as we think of these when asking whether someone else was justified in doing what she did.

¹¹ The fact that *p* is false may well explain *q*, but *the fact* that *p* is false is, itself, a truth.

¹² Thus, Parfit, Dancy, and Schroeder’s use of ‘motivating,’ unfortunately, differ from Smith’s. (Thanks to Nicholas Silins, on this point.) What the others call ‘motivating reasons,’ Smith often calls ‘my normative reason.’

facts about her beliefs to count in favor of leaving (as she might if, say, all who did not share the beliefs of the congregation were asked to leave). Rather, she took the meeting's end to count in favor of leaving. And, though she was wrong about the meeting's end, she was right to take its end, rather than her belief about its end, to be what counted in favor of leaving. Only so can we say that, since the meeting was not over, there was nothing that counted in favor of leaving, even though Erin thought there was. And only so can we say that, if the meeting is in fact over, that fact counts in favor of leaving even if Erin is unaware of it. If we were instead to insist that our beliefs, themselves, are what count in favor of acting, we would have to say that we do not, by making our beliefs more accurate, thereby improve our information about what we have reason to do. This is unacceptable.¹³

So, people typically take facts about the world at large, rather than facts about their own psychology, to be what counts in favor of action, and they are typically correct to do so. And yet, in light of the unfortunate possibility of error, it can seem that one's own (operative) reason for action cannot explain one's action, at least in the case of error. Thus it seems that the reasons that explain action and the agent's own reasons for acting are different kinds of things. Some find this worrying, as we will see.

First, though, to finish our taxonomy: We now have, on stage, three contenders for the title 'reasons for action:' first, considerations that (in fact, truly) count in favor of acting, which Scanlon calls 'reasons in the standard, normative sense' and others sometimes call 'normative reasons;' second, considerations that someone *took* to count in favor of acting, on the basis

¹³ Even Bernard Williams, who argues from the assumption that reasons for action must be able to explain the action for which they are reasons to some significant restrictions on reasons for action, does not insist that your beliefs are themselves either what counts in favor of action or what you take to count in favor of acting. Even on his view, you have a reason not to drink the petrol, and no reason to do so, even when you believe that it is gin and desire to drink a gin and tonic. (Williams insists you have a reason only if it *possible*, given certain idealizations, for you to believe that you have that reason. See Williams, 'Internal and External Reasons'.) So, it is a quite extreme, and I think implausible, to think that beliefs themselves are what counts in favor of acting. (Mark Schroeder avoids the implausible conclusion by allowing the required mental state to be a 'background condition.' See Schroeder, *Slaves of the Passions* at 23-40.).

of which he or she acted (considerations someone treated as reasons in the standard normative sense), which Scanlon calls ‘operative reasons’ and which are sometimes referred to as ‘the agent’s reasons’ (and often as ‘motivating reasons’); and, finally, considerations that explain an action, whatever these may be. Given the possibility of error, this last role seems to be played by psychological states of the agent, which Smith calls ‘motivating reasons.’ By keeping in mind that there are (at least) these three possible characters, we can minimize confusion.¹⁴

One might then be a pluralist about the phrase ‘reasons for actions,’ allowing the title to all three contenders. I would recommend such pluralism, believing it acceptable to all, so long as we do not insist that the reasons that explain action are psychological states. Some philosophers, such as Dancy and Frederick Stoutland, deny this (as we will see). So, to claim that reasons that explain action are psychological states is to go beyond taxonomy and adopt a disputable view. But even Dancy and Stoutland can allow that there are (at least) three different roles a state of affairs, event, or consideration can play, each of which is in some way deserving of the title ‘reason for that action:’ it could (genuinely) count in favor of the action, it could be taken (correctly or incorrectly) to count in favor of the action by someone who acted on it, and it could explain the action. To deny the title ‘reason for action’ to any of these three contenders is to ask for confusion.

¹⁴ For example, Dancy, in *Practical Reality*, appeals to what he calls the ‘normative constraint’ in arguing that the agent’s own reasons for action are what explain her action: he insists that ‘normative reasons’ and ‘motivating reasons’ must be ‘at least capable of being the same kind of thing.’ By ‘normative reasons’ Dancy means something like considerations that in fact count in favor of acting (what we are calling ‘reasons in the standard normative sense’). But it is sometimes unclear whether Dancy means, by ‘motivating reasons,’ what we are calling ‘operative reasons’ (that which the person took to count in favor of acting), or, instead, if he means to refer to whatever it is that explains the action. Understood the first way, the constraint is easily accepted by all, but it does not, then, seem to me to do the work he hopes it do. Understood the second way, it seems to assume the position for which Dancy hopes to argue. One can worry that the argument works simply by claiming that, because *operative* reasons must be ‘at least be capable of being the same kind of thing as’ reasons in the standard normative sense, so, too, must that which explains the action. In the symposium on the book, both Stephen Darwall and Wayne A. Davis suggest that this issue needs to be clarified. See Stephen Darwall, ‘Desires, Reasons, and Causes’, *Philosophy and Phenomenological Research*, 67/2 (2003), 436-43, Wayne A. Davis, ‘Psychologism and Humeanism’, *Ibid.* (452-59).

2. EXPLAINING ACTION

Having accomplished the taxonomy, let us consider, in more detail, reasons that play the third role: how do we explain action? In particular, remembering Davidson's opening question, how do we explain action *by* giving the agent's own, operative, reason?

We have already considered the very simple position, according to which the agent's own (operative) reason itself explains the action, and we have raised an important complaint against it: since the agent's own reasons are sometimes false, this view would insist that falsehoods (not the fact that something is false, but the falsehood itself) sometimes explain what happens.

Surprisingly, Dancy accepts this consequence and occupies the simple position. According to Dancy, the reason that explains Erin's departure is *that the meeting was over* (not that *Erin believed* that the meeting was over).¹⁵ Dancy acknowledges that his position requires him to deny what he refers to as 'the simple thought that all explanation is factive.'¹⁶ Given this steep cost, I want to consider why one would be drawn to this kind of position.

The appeal, for Dancy, lies in avoiding the apparently worrisome gap we noted earlier: If psychological states explain action, then the reasons that explain the action and the agent's own reasons for acting are starkly different kinds of things.¹⁷ Erin's action is to be explained by facts about her psychology, while, from her point of view, that which counts in favor of acting has nothing to do with her psychology. This gap—which I will call 'Dancy's gap'—can seem unsettling. It might seem that our operative reasons should play a crucial role in explaining our action. However, if, when Erin is mistaken, her departure is explained by facts

¹⁵ See, especially, Dancy, *Practical Reality* at 135ff. He says, 'One contentious aspect of the picture that has been developed here is that something that is not the case can explain an action' Ibid. at 137. He does allow facts about the agent's psychology to be what he calls 'enabling conditions.'

¹⁶ Ibid. at 131.

¹⁷ I believe this gap is the target of Dancy's 'normative constraint.' Many philosophers are unbothered it.

about her psychology, it can seem that those same facts will explain her departure, even when she is correct. The meeting's end can seem dispensable.¹⁸ In contrast, Dancy's view preserves a crucial role for it.

Thomas Nagel voices a different worry about explaining action by appeal to psychological states:

The recognition of reasons as reasons [in the standard normative sense] is to be contrasted with their use purely as a form of psychological explanation. [Nagel here cites Davidson's article.] The latter merely connects action with the agent's desires and beliefs, without touching the normative question of whether he *had* any adequate reason for acting—whether he should have acted as he did. If this is all that can be said once we leave the point of view of the agent behind, then I think it would follow that we don't really act for reasons at all. Rather, we are caused to act by desires and beliefs, and the terminology of reason can be used only in a diminished sense to express this kind of explanation. (Nagel 1986: 142)

Nagel here suggests that, if the reasons that explain action bear no clear relation to that which might count in favor of action—if, to use Nagel's term, the explanation of action does not 'touch' the question of whether there was adequate reason (in the standard, normative sense) to act—then it can seem that 'we don't really act for reasons at all.'

While Dancy's gap stands between the agent's own (operative) reasons and the psychology that purportedly explains action (the motivating reasons), Nagel draws our attention to a different gap—between that which explains an action and reasons in the standard normative sense. He suggests that the gap between these cannot simply be left unanalyzed, in our final story of the explanation of actions done for reasons. This seems to me correct. Our explanation of action done for reasons should do something to relate that which explains the action to reasons in the standard, normative sense.

Dancy's account provides such a relation. According to Dancy, the reasons that explain action are the operative reasons, i.e., the considerations the agent took to count in favor of

¹⁸ Dancy's worries echo certain forms of skepticism: once we appeal to psychology, or appearances, to account for the non-veridical cases of belief, it can seem hard to recover an essential role for the outside world, in the veridical ones. I will not here explore the parallels, nor the possibility of a 'disjunctive' solution. I only mean to provide some motivation for Dancy's worry. Though I see the worry, I am not, myself, gripped by it.

acting, on the basis of which she acted. Thus, the reasons that explain the action were made to ‘touch the question of whether there was adequate reason to act’ by the agent herself. By taking them to count in favor of acting and acting upon them, she *treated them* as adequate reason to act.¹⁹ So Nagel’s worry is stilled: it is clear that we do ‘really’ act for reasons—in particular, for what we treat as reasons in the standard normative sense.

Despite its success in avoiding these worries, Dancy’s view still seems unacceptable: you cannot explain what happened by appeal to a falsehood. A weaker view (which perhaps still captures Dancy’s thought, and which, I believe, captures the spirit of the theories that were the original targets of Davidson’s article) would instead deny that the activities of agents, *qua* rational activities, can be properly explained—or, less tendentiously, that they are to be explained in the usual way in which one explains other (mere) happenings.

According to such a view, the two questions ‘Why did Erin leave?’ and ‘Why did the engine fail?’ bear only surface similarity. If you ask ‘Why did the engine fail?’ you are asking what is sometimes called a theoretical question, asking, in effect, ‘How did it come about that the engine failed?’ in a quasi-scientific spirit. By answering such a question, you provide an ordinary explanation. But, one might think, one does not ask *this* sort of question, when trying to ‘explain’—to make intelligible—action. Rather, when asking, ‘Why did Erin leave?’ one is in fact asking a markedly different question, *viz.*, ‘From her point of view, why leave?’ To ‘explain’ an action, *qua* action, then, is not to answer a theoretical, quasi-scientific question about an ordinary event; it is not to say *how it came about* that Erin left. It is rather to answer a *deliberative* question, but now framed from the agent’s point of view: it is to say what, from her point of view, counted in favor of leaving. Note that, by explaining the action in this way, we address Nagel’s concern: we preserve a connection between the action

¹⁹ As we will see, the same cannot be said of the contents of a belief-desire pair that causes the action: they were not necessarily treated by the agent as considerations that count in favor of acting.

and the question of whether the action was called for. The reasons we appeal to, in explaining the action, are the reasons the agent might use, in deciding whether to leave. It will, of course, be entirely unremarkable that such ‘explanations,’ framed as they are from another’s point of view, sometimes refer to falsehoods. When they do, then, to avoid confusion, we will mark that fact by saying, ‘Erin left the room because *she thought* the meeting was over.’ But, in this context, the addition of ‘she thought’ does not contribute to the explanation. It simply makes explicit what is true in any such explanation: it is given from the agent’s point of view.²⁰

While I have great deal of sympathy for the view that, when explaining action, we are engaged in a quite different enterprise than when explaining other events, and while it seems to me that, in a many cases in ordinary life, when we ask ‘why did so-and-so do such-and-such?’ we are most interested in, and perhaps ‘really’ asking, ‘from so-and-so’s point of view, why do so such and such?’, it was a position of this sort that Davidson’s article displaced.

Davidson, in effect, pointed out that there may be a great many possible answers to the question, ‘From her point of view, why leave?’ which played no role in her leaving (because, e.g., she did not notice them). In answering the question ‘from her point of view, why do thus-and-such?’, and so making intelligible why someone could or would or might so act,

²⁰ That Dancy ascribes to such a view is suggested by his claims: ‘We explain the action by showing that the answer to the . . . question [Had things been the way he supposed them to be, would his action have been the one there was most reason to do?] is yes. . . . to explain an action is to justify it only in a certain sense’ Ibid. at 9. Later he says, ‘The explanation of an action succeeds to the extent that it enables us to see how the agent might have taken certain features of the action as good reasons to do it’ Ibid. at 95. (Note the “might have.”)

Stoutland suggests (in conversation) that these two forms of explanation are joined by the thought that they are making something intelligible. One makes engine failures intelligible by noting their cause, but makes actions intelligible by giving the agent’s (operative) reasons.

Whether activities can be explained ‘in the way in which one explains mere happenings like engine failure’ should, I think, be distinguished from the question of whether that explanation is ‘causal’ in any restrictive sense.

Finally, note that this line of thought need not deny that there are other ways of explaining the event of her departure—neural explanations, for example. See Ibid. at 176-77.

one has simply noted relations of justification holding between features of the situation. One has not, yet, done anything to explain what actually happened.

As I understand it, Davidson's article displaced the earlier view simply by reasserting a demand for a more ordinary form of explanation: an explanation which shows, not merely what, from another's point of view, *could* count in favor of acting, but why that person did, in fact, act. This demand seems to me appropriate. Moreover, because this explanation is framed from our own point of view, it must appeal to what are (from our point of view) facts. (Though the demand seems to fall short of insisting on a causal relation—unless one thinks any explanation of a happening is a species of causal explanation.)

We have, then, arrived at some criteria. We would like to satisfy Davidson's demand—to explain what actually happened, by appeal to truths. We will not, then, close Dancy's gap. Nonetheless, in light of the worries expressed by Dancy and Nagel, the explanation should preserve some clear relation between that which explains the action and the agent's own, operative reasons—that is, those considerations she treated as reasons in the standard, normative sense. We want to explain what in fact happened by appeal to the agent's operative reasons. We seem to have returned to the neighborhood of Davidson's opening question.

Davidson's own position might seem, at first, to satisfy. He claims that we explain how it came about that Erin left by appeal to a belief-desire pair.²¹ He also claims that the belief-desire pair explains the action *by* giving the agent's reason for doing what she did. (The pair thereby *rationalizes* the action, in Davidson's terms.) How does the belief-desire pair give the agent's reason? It does so insofar as we can recover, from the content of the belief and

²¹ More exactly, the position for which he argues is that 'For us to understand how a reason of any kind rationalizes an action [i.e., explains it by providing the agent's reason for it] it is necessary and sufficient that we see, at least in essential outline, how to construct a primary reason.' A primary reason is a belief-desire pair. Davidson, 'Actions, Reasons, and Causes', at 4.

desire, reference to ‘some feature, consequence, or aspect of the action the agent wanted, desired, held dear, thought dutiful, beneficial, obligatory, or agreeable.’ These features, consequences, or aspects of the action seem to be that which the agent took to count in favor of acting. It may seem, then, that Davidson has explained an action done for a reason by appeal to facts, while preserving a clear role for the agent’s own (operative) reasons.

As time went on, however, Davidson came to see that the role he had provided for the agent’s reasons was insufficient. To illustrate, he provides an example in which a climber lets go of a rope because his belief that doing so would make him safer and his desire to be safer so ‘so unnerve him as to cause him to loosen his hold.’ Yet, this climber ‘never chose to loosen his hold, nor did he do it intentionally.’ Davidson continues,

Beliefs and desires that would rationalize an action if they cause it in the *right* way—through a course of practical reasoning, as we might try saying—may cause it in other ways. If so, the action was not performed with the intention that we could have read off from the attitudes that caused it. What I despair of spelling out is the way in which attitudes must cause actions if they are to rationalize the action. (Davidson 1980b: 79)

This example is typically taken to illustrate the possibility of so-called ‘deviant causal chains’ and so to call for some specification the ‘right’ causal relation. I want to set this issue aside, and focus, instead, on how the climber example illustrates what I have called Nagel’s gap (and Davidson’s sensitivity to it).²²

Davidson criticized his opponents for merely noting that relations of justification hold between features of a situation, without explaining how or why anything actually happened. Davidson tried to answer his criticism by, in effect, pairing relations of justification with

²² My position, in this paper, leaves open the possibility of finding the appropriate causal relations. I suspect, though, that Davidson’s despair was not born simply of the difficulty or complexity of the task. Rather, his anomalous monism suggests an in-principle barrier to specifying the right causal relation. He later says, ‘I do not see how the right sort of causal process can be distinguished without, among other things, giving an account of how a decision is reached in the light of conflicting evidence and conflicting desires. I doubt whether it is possible to provide such an account at all, but certainly it cannot be done without using notions like evidence, or good reason for believing, and these notions outrun those with which we began.’ Donald Davidson, ‘Philosophy as Psychology’, *Essays on Action and Events* (Oxford: Oxford University Press, 1980a), 229-44 at 232–33.

causal relations: he dropped the justifying considerations into mental states that (via their relation to brain events) causally explain actions.²³ His account does, at least, explain why something happened.

However, the climber example illustrates that simply dropping certain contents into mental states that, as a matter of fact, provide a causal explanation does not, yet, close Nagel's gap. Even though the contents of the relevant mental states may (from our point of view, or even from the agent's own point of view) bear the right justificatory relation to the event, it is not yet clear that those contents were *treated*, by the agent, as reasons in the standard normative sense, nor, crucially, that the agent's so treating them has any role to play in the explanation of what, in fact, happened—it is not clear that those contents played the role of anyone's operative reasons. Our explanation has not yet captured the appropriate role for the agent's own reasons. And so it may still be that, 'we don't really act for reasons at all. Rather, we are caused to act by desires and beliefs.' Davidson recognizes that we want the agent's own reasons to take their place in something like 'a course of practical reasoning.' Further, we want their playing such a role to be part of our explanation of what actually happened.²⁴

So Davidson's original account does not satisfy our criteria. To explain an event by appeal to a causal relation between that event and a belief-desire pair whose contents, as it happens, also make the action intelligible, is just to provide an explanation that includes

²³ More strictly: potentially justifying considerations are the content of mental states whose onset is token-identical with brain events standing in causal relations to the bodily events which instantiate actions.

²⁴ J. David Velleman raises a related worry with Davidson's account: he wonders whether the role of the mechanisms specified in Davidson's account add up to the *agent's* role in producing action. He suggests they do not. He suggests a mechanism that he believes plays the agent's role. See J. David Velleman, 'Introduction', *The Possibility of Practical Reason* (Oxford: Oxford University Press, 2000). I would understand his account, not as a competitor to the account I am about to offer, but rather as one way of filling it in.

Jason Dickenson has argued that Davidson's form of explanation does not meet his own central challenge for a different, though related, reason: he argues that the explanation in question is what he calls 'contrastive' and that Davidson's answer cannot explain why an agent acted for only one of two reasons which she recognizes. See Jason Dickenson, 'Reasons, Causes, and Contrasts', *Pacific Philosophical Quarterly*, 88/1 (March 2007), 1-23.

reference to considerations for which someone *could* act. We want an explanation that shows why or how someone *in fact* acted for these (rather than those) reasons. We want to better understand how the agent's own reasons for acting play a distinctive role in explaining what in fact happened.²⁵

3. AN ALTERNATIVE FORM FOR ACTION EXPLANATION

We remain, then, very much in the dark. I suggest that, given our current state of disagreement and confusion, it is worth specifying a general form or pattern of action explanation to which all might agree. I am therefore emboldened to propose an extraordinarily minimal, embarrassingly simple, and yet, I believe, largely satisfying account, one that promises both to explain the action by appeal to what are, from the explainer's point of view, facts and to make clear that the event was an action done for certain reasons (if it was done for reasons²⁶). Minimal and formal as it is, there is much that this simple account does not do. In fact, some will think it simply a kind of analysis, and that before the analysis is explanatory, its pieces must, themselves, be explained. This may be. Nonetheless, given the difficulties so far encountered, I believe providing such an analysis, or specifying the form the explanation should take, can be of use.

The proposal starts with this simple thought: whenever an agent acts for reasons, the agent, in some sense, takes certain considerations to settle the question of whether so to act, therein intends so to act, and executes that intention in action.

If this much is uncontroversial (and, under some interpretation, I believe it must be), we can use it as a form for filling out. I propose, then, that we explain an event that is an action

²⁵ Davidson himself doubts that the role of reasons in practical deliberation can be captured in causal terms. See Davidson, 'Philosophy as Psychology', at 232.

²⁶ The account accommodates action for no (particular) reason (by allowing that we can settle a question for no reason).

done for reasons by appealing to the fact that the agent took certain considerations to settle the question of whether to act in some way, therein intended so to act, and successfully executed that intention in action. I suggest that *this* complex fact, rather than a belief-desire pair, is the reason that rationalizes the action—that explains the action by giving the agent’s reason for acting.

The proposal’s main advantages derive from its underlying understanding of what a reason is. While most others understand reasons in the standard normative sense as we have been **understanding them thus** far—as considerations that *count in favor of* actions and attitudes (or, equally, considerations that *show something good* [important, desirable, etc.] about actions or attitudes)—I have argued elsewhere that understanding reasons as considerations standing in relation to actions or attitudes (via the *counting in favor of* or *showing good* relation) generates an ambiguity, known as the wrong kind of reasons problem.²⁷ I have suggested that we instead understand reasons as items in pieces of reasoning, where reasoning is thought (or possible thought) directed toward some conclusion. Reasons, then, are considerations that bear or are taken to bear on questions. Adopting this account, we can no longer relate reasons directly to actions or attitudes. Rather, we are forced to relate reasons to actions or attitudes *via* the question on which they (are taken to) bear. Forcing ourselves to consider the question on which a reason bears or is taken to bear not only allows us to solve the wrong kind of reasons problem, but also has other, more momentous, advantages, some of which are on display in the explanation of action.²⁸

²⁷ See Hieronymi, 'The Wrong Kind of Reason', (The ambiguity appears in the case of certain attitudes, such as belief or intention, but it does not appear in the case of ordinary, voluntary action. Since we have been discussing only action, our discussion, thus far, has not been plagued by the ambiguity.

²⁸ Becoming grand and gestural, for a moment: relating considerations directly to actions or attitudes—to events or psychology—seems to me not kosher, a kind of unholy blending of the rational and the empirical. As I will suggest below, so relating them occludes just the thing we are after, in explaining action: the role of the agent. She is the one who settles questions and therein acts (or intends, or believes...). Her activity is what, so to speak, mediates between the rational and the empirical.

First, and simply, if reasons are, fundamentally, considerations that (are taken to) bear on questions, we need not divide reasons at the root, between ‘explanatory’ and ‘normative.’ Rather, we distinguish different kinds of reasons according to the different kinds of questions on which they bear. We face (implicitly or explicitly) the question of whether to act. Reasons, in the standard normative sense, for a given action are facts that bear positively on the question of whether so to act.²⁹ We also ask and answer questions about why things happen or why things are as they are. The considerations that answer such questions explain the relevant happening or state of affairs. These will be ‘explanatory’ reasons.³⁰

Reconsider the hybrid case: explaining events that are themselves done for reasons, such as Erin’s leaving the room. We can now answer our explanatory question (How did it come about that Erin left?) by appeal to the fact that Erin settled some *other* question (whether to leave). In answering our explanatory question, we will appeal to the fact that Erin settled *her* practical question, for her (operative) reason. Her operative reason—the consideration she treated as a reason in the standard normative sense, in answering her question—appears among our explanatory reasons; but it appears *as* her operative reason, bearing, for her, on her question. ~~By making reference to her question, we thus answer our explanatory question while preserving the proper role for her operative reason, by making reference to her question.~~

Notice we thereby close Nagel’s gap. The account provides a fairly clear view of the relation between reasons in the standard normative sense and the reason that explains the

²⁹ We also face (implicitly or explicitly) the question of whether this or that is so. Facts that bear positively on such a question will be reasons, in the standard normative sense, for believing this or that. Considerations that someone took to settle the question of whether this or that is so will be that person’s own (operative) reasons for believing this or that—the reasons on which her belief is ‘based,’ to use the going terminology.

³⁰ This way of understanding reasons and attitudes makes the notion of a ‘normative reason’ seem strange. All reasons bear on questions, and the questions on which the so-called ‘normative’ reasons for belief, intention, or action bear are not especially normative. ‘Normative’ reasons for believing the recession is over are considerations that bear on the question of whether the recession is over. These do not seem any more or less ‘normative’ than the ‘explanatory’ reasons that answer the question of why the recession ended.

action: the complicated fact that explains the action includes within it the fact that the agent treated certain considerations as reasons in the standard normative sense—she took them to bear on (in fact, to settle) the question of whether to act. It is thus clear that we do (sometimes) act on reasons. Following Davidson’s intuitions, we closed this gap by providing ourselves with something like ‘a course of practical reasoning’ (albeit a very short one).³¹

We also avoid Davidson’s criticisms in ‘Actions, Reasons, and Causes.’ In answering our explanatory question (How did it come about that Erin left?), we have done more than make the action intelligible from Erin’s point of view. We have claimed, more robustly, that certain considerations, and not others, were those for which the Erin, in fact, formed an intention, which intention she executed in the event that was the action. We have satisfied the demand for a more ordinary form of explanation.

However, it is not clear that the relation between the complex fact that explains by giving the agent’s reasons and the event that is the action is causal one (unless one assumes that all explanations of happenings are causal). The account is agnostic, on this question. So it may remain acceptable to those, like Stoutland, who resist explaining actions by appeal to causes.

Relatedly, the account avoids the possibility of so-called ‘deviant causal chains.’ According to the proposed account, the agent, for certain reasons, settles the question of whether to act, *therein* intends to act, and executes *that* intention in the event that is the action. The connections are too tight for deviance.

In light of these last observations, some will conclude that the complex fact does not actually explain the action. Rather, it provides a kind of analysis of action done for reasons, one which will become explanatory only as its pieces are, in turn, explained. This claim raises interesting questions that I will not take up. I will be content if it is agreed that, if one is to

³¹ Velleman’s worry need not arise, either: it is I, the agent, who settles questions and therein intends.

explain action in a way that preserves the role of the agent's reason for acting, one's explanation should be able to be fit into the form or analysis I here propose.³²

Although the proposal does not itself provide a causal explanation of Erin's leaving, it also does not rule out such an explanation. Rather, it provides a basic form into which any (more elaborate) explanation should endeavor to fit. Thus, more ambitious theories should be able to be seen, not as competitors, but as filling in the proposed account. Perhaps we will locate the mental states, events, and causal relations (or, the neurological and biological goings-on) whose instantiation amounts to the agent taking some purported fact to settle the question of whether to act in some way, therein intending so to act, and executing that intention in action. We would then have found the 'right' causal chain from the agent's operative reason to the action. Whether there are (or could be) independently identifiable causal relations that *generally* underlie the complex fact across different instances of it seems an interesting question about emergence and reduction, which the current proposal does not address.

³² Approaching the interesting questions, by analogy: Suppose you ask, 'How did it come about that Erin won the match?' and I answer, 'She earned more points than her opponent.' My answer may be (annoying, because) uninformative. (It would be informative only if there is some other way for Erin to win—perhaps by knock-out.) Nonetheless, it seems a correct answer to the question: she did, in fact, win by earning more points. (In fact, she could not have won in any other way, and this is known to be so.) Perhaps the answer would become informative and so explanatory if I add some detail: she earned seven points to her opponent's five? In any case, if we care to provide some (further) explanation of her winning, our explanation must explain how she earned more points. Otherwise we will have explained, not her winning, but some other, nearby fact (e.g., how she repeatedly brought the racquet into contact with the ball). The uninformative answer, then, provides the form that any more informative explanation must take.

Likewise, if we ask how it came about that Erin left the room, and we already know that her leaving was an action done for reasons, it may be uninformative to be told that she left because she took certain considerations to settle the question of whether to leave, therein intended to leave, and executed that intention. Nonetheless, it seems correct as answer: that was, in fact, how it came about that Erin left. Perhaps the answer will become informative and explanatory if I add *which* considerations she took to settle the question. In any case, if we want to provide some further explanation of her leaving—where what we want to explain is her action, and not some other nearby fact (e.g., how her body came to be at a certain place at a certain time)—our further explanation must explain, in turn, how it was that Erin took those considerations to settle that question, therein intended, and executed the intention in action.

An underlying question seems to be, if a certain kind of phenomena is, and is known to be, the result of (or constituted by) a certain sort of process, does citing the fact that such a process occurred explain an instance of that phenomena? I do not think this question needs to be resolved for the purposes at hand.

Dancy's gap remains: we explain action by appeal to certain broadly psychological activities. It may thus still seem that that which the agent took to count in favor of leaving—the meeting's end, in Erin's case—is dispensable, at least in the case of error. This seems unavoidable. However, once we have addressed Nagel's worry, I believe this remaining gap need not trouble us. It seems appropriate that, in the explanation of her action, the agent's activities should, so to speak, 'stand in' for those (purported) facts that she takes to be reason-giving. Her taking them to be reasons explains her action. Thus, she is accountable for her actions—she, not the facts that call for action, brings her action to be.

So, I think even this minimal, formal account accomplishes quite a bit. However, there is a great deal the proposal does not accomplish. It makes no attempt to explain what it is to 'take' a reason to settle the question of whether to act, and therein intend, for that reason (taking c to settle the question of whether to act is not believing that c is a conclusive reason to act; it is rather treating c as a conclusive reason to act). Again, one might try to give an account of the mechanisms underlying such activities.³³ The proposal makes no attempt to explain why an agent takes a given consideration to count in favor of an action or how she does so. Nor does it make any attempt to explain the relation between intending and acting, or between mind and movement. These are huge, remaining gaps.

Nonetheless, we can now see that the most momentous advantage of changing the underlying account of what a reason is lies not simply in providing a genus into which species of reason can fit, nor even in avoiding the wrong kind of reasons problem, but rather in allowing us access to the activity of the agent, as she responds to reasons. By exposing that activity, we can move from merely noting relations of justification—which we

³³ As I understand it, this is Velleman's explicit ambition.

might pair with causal relations—to explaining what actually happened by appeal to the agent’s own reasons for acting.

4. CONCLUSION

I have offered a suggestion about how to relate the reasons that explain an action to the agent’s own reasons for acting. I believe it occupies a middle ground between Davidson, in ‘Actions, Reasons, and Causes,’ and his opponents. Davidson opens that article by asking, ‘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?’ He suggests that the reason that explains the action by giving the agent’s reason for doing what he did is a belief-desire pair. He believes that the relation between this reason and the action must be causal, to enable us to pick out, from among the many reasons that a person might have acted upon, those that the person did, in fact, act upon. I have suggested, instead, that the fact that explains the action by giving the agent’s reason for doing what he did is the fact that the agent took certain considerations to settle the question of whether to act, therein intended to act, and executed that intention in action. This account avoids both Davidson’s criticisms and the difficulties that have arisen for Davidson’s account and others. In being so very minimal, my proposal may be a step backwards. Nevertheless, I hope it is a step onto solid ground, from which we might find a new and more fruitful direction to pursue.³⁴

³⁴ Versions of this material benefited, over the years, from conversation or commentary from Tyler Burge, Sarah Buss, John Carriero, Jonathan Dancy, Sonny Elizondo, Mark Greenberg, Mark C. Johnson, Sean Kelsey, Niko Kolodny, Aaron James, Matt McAdam, Simon Rippon, T. M. Scanlon, Mark Schroeder, Sheldon R. Smith, Nicolas Southwood, Frederick Stoutland, Julie Tannenbaum, and Gideon Yaffe, as well as members of my graduate seminar at UCLA and audiences at the University of Sydney, the University of Texas at Austin, Texas A&M, Southern Methodist University, the University of Miami, the Aristotelian Society, and Oxford University.

REFERENCES

- Broome, John (2004), 'Reasons', in R. Jay Wallace, et al. (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (Oxford: Clarendon Press), 28–55.
- Dancy, Jonathan (2000), *Practical Reality* (Oxford: Oxford University Press) 187.
- Darwall, Stephen (2003), 'Desires, Reasons, and Causes', *Philosophy and Phenomenological Research*, 67 (2), 436-43.
- Davidson, Donald (1980a), 'Philosophy as Psychology', *Essays on Action and Events* (Oxford: Oxford University Press), 229-44.
- (1980b), 'Freedom to Act', *Essays on Actions and Events* (Oxford: Oxford University Press), 63_81.
- (1980c), 'Actions, Reasons, and Causes', *Essays on Actions and Events* (New York: Oxford University Press), 3–19.
- Davis, Wayne A. (2003), 'Psychologism and Humeanism', *Philosophy and Phenomenological Research*, 67 (2), 452-59.
- Dickenson, Jason (2007), 'Reasons, Causes, and Contrasts', *Pacific Philosophical Quarterly*, 88 (1), 1-23.
- Gibbard, Alan (1990), *Wise Choices, Apt Feelings* (Cambridge, MA: Harvard University Press).
- Hieronymi, Pamela (2005), 'The Wrong Kind of Reason', *The Journal of Philosophy*, 102 (9), 1–21.
- Nagel, Thomas (1986), *The View From Nowhere* (Oxford: Oxford University Press).
- Parfit, Derek (2001), 'Rationality and Reasons', in Dan Egonsson, et al. (eds.), *Exploring Practical Philosophy* (Burlington, VT: Ashgate Press), 17–39.
- Raz, Joseph (1975), *Practical Reason and Norms* (London: Hutchinson).
- (1999), *Engaging Reason: On the Theory of Value and Action* (Oxford: Oxford University Press).
- Scanlon, T. M. (1998), *What We Owe to Each Other* (Cambridge, MA: Harvard University Press).
- Schroeder, Mark (2008), *Slaves of the Passions* (Oxford: Oxford University Press) 224.
- Skorupski, John (1997), 'Reasons and Reason', in Garrett Cullity and Berys Gaut (eds.), *Ethics and Practical Reason* (Oxford: Oxford University Press), 345–67.
- Smith, Michael (1994), *The Moral Problem* (Oxford: Blackwell).
- Velleman, J. David (2000), 'Introduction', *The Possibility of Practical Reason* (Oxford: Oxford University Press).
- Williams, Bernard (1981), 'Internal and external reasons', *Moral luck*, 101-13.