

# INSTRUMENTAL DESIRES, INSTRUMENTAL RATIONALITY

by Michael Smith and Edward Harcourt

*I—Michael Smith*

**ABSTRACT** The requirements of instrumental rationality are often thought to be normative conditions on choice or intention, but this is a mistake. Instrumental rationality is best understood as a requirement of coherence on an agent's non-instrumental desires and means-end beliefs. Since only a subset of an agent's means-end beliefs concern possible actions, the connection with intention is thus more oblique. This requirement of coherence can be satisfied either locally or more globally, it may be only one among a number of such requirements on an agent's total set of desires and beliefs, and it has no special connection with reasoning. An appreciation of these facts leads to a better understanding of both the nature and the significance of instrumental rationality.

A number of theorists have recently attempted to account for the normativity of instrumental rationality.<sup>1</sup> Though there are significant differences, their explanations all share a common feature: each assumes that the requirements of instrumental rationality are normative conditions, of some sort, on choice or intention. If instrumental rationality were limited to situations in which we make choices or form intentions then these explanations would have some chance of being correct. But instrumental rationality is not limited in this way.

My aim in the present paper is thus to lay out an alternative explanation, one that accords instrumental rationality its proper scope. I proceed by spelling out the alternative in the form of a series of assertions which I expand upon and justify. I do not claim any originality, as the ideas on offer will be familiar from standard accounts of decision theory. My reason for putting

1. Christine Korsgaard, 'The Normativity of Instrumental Reason' in Garrett Cullity and Berys Gaut (eds.), *Ethics and Practical Reason* (Oxford: Oxford University Press, 1997); John Broome, 'Normative requirements', *Ratio* 12 (1999), 398–419; R. Jay Wallace, 'Normativity, Commitment, and Instrumental Reason' in *Philosophers' Imprint*, <[www.philosophersimprint.org/001003](http://www.philosophersimprint.org/001003)> 1 (2001), 1–26.

them forward, nonetheless, is to provide the needed reminder of these ideas' significance in the context of recent discussions.<sup>2</sup>

## I

*Agents typically have both instrumental and non-instrumental desires.*

Consider the following passage from Hume.<sup>3</sup>

Ask a man *why he uses exercise*; he will answer *because he desires to keep his health*. If you then enquire, *why he desires to keep his health*, he will readily reply, *because sickness is painful*. If you push your enquiries farther, and desire a reason *why he hates pain*, it is impossible he can ever give any. This is an ultimate end, and is never referred to any other object.

Hume's idea here is that, much as our actions are explicable in terms of our desires and beliefs, so certain of our desires are susceptible to that same sort of explanation. We can explain why we desire some of the things we desire by citing other desires we have and beliefs about how, if the first of the things we desire comes about, this will lead to the other thing we desire coming about.

In the simplest case one thing we desire leads to another by causing it. For example, we explain why I desire to exercise by citing the fact that I desire health and believe that exercise causes health. But there are other ways in which one thing can lead to another as well. To use a familiar example, I may desire to raise

2. As will become clear to those familiar with their work, while my disagreement with Broome, and perhaps with Wallace too, turns out to be mainly a disagreement about the scope of instrumental rationality, my disagreement with Korsgaard is more complete, as she thinks that the requirements of instrumental rationality can be restated as claims about what we have reason to do. (Though note that Wallace sometimes says this sort of thing too: see, for example, his 'Practical Reason' in *The Stanford Encyclopedia of Philosophy* (Winter 2003 Edition), Edward N. Zalta (ed.), forthcoming URL = <<http://plato.stanford.edu/archives/win2003/entries/practical-reason/>> .) But this seems to me quite wrong, as it would require a connection between instrumental rationality and instrumental reasoning that does not exist. I will not, however, offer any explicit criticisms of these authors beyond what I have already said. My aim, to repeat, is to spell out an alternative. For more on this see my 'Is there a Nexus between Reasons and Rationality?' forthcoming in *Poznan Studies in the Philosophy of Science and Humanities: New Trends in Moral Psychology* (ed.) Sergio Tenenbaum (Amsterdam: Rodophi).

3. David Hume, *Enquiries concerning Human Understanding and concerning the Principles of Morals*, 1777 edition, (Oxford: Clarendon Press, 1975), Appendix I.

my arm because I desire to signal a turn and believe that raising my arm leads to signaling a turn, not in the sense that raising an arm causes a signaling, but rather in the sense that, given the conventions hereabouts, raising an arm constitutes giving a signal. In what follows I will call all beliefs about the ways in which one thing leads to another 'means-end beliefs'. Though the means-end beliefs discussed will typically be causal beliefs, we should bear in mind the fact that this is a simplification.

The main point Hume makes in the passage is that we cannot explain every desire we have by citing another desire and a means-end belief. Such chains of explanation must come to an end somewhere. As he sees things a chain of explanation like that described would come to an end with the desire to avoid pain. But this is, of course, an empirical claim, one which we may or may not think is typically true. People can desire anything. Moreover no particular empirical claim is crucial. What is crucial is rather the suggestion that such chains of explanation must come to an end with some desire or other which eludes explanation in this way. In some people, some of the time, the chain might come to an end with the desire to avoid pain, but in others it might come to an end with the desire for health, and in some people, some of the time, it might come to an end with the desire to exercise, or even with the desire to move their body.

Though Hume calls the desires that come at the origin of such chains of explanation 'ultimate' desires, I will call them 'non-instrumental' desires, and, accordingly, I will call the desires that are explained by such desires and means-end beliefs 'instrumental' desires. My reason for baulking at the label 'ultimate' is that this seems to suggest, misleadingly, that no rational explanation can be given of why we have the non-instrumental desires we have. But though this was Hume's own view, it seems to me quite wrong.<sup>4</sup> I may have a non-instrumental desire that p because I have non-instrumental desires that q, r, and s, where the desires that p, q, r, and s all fit together in a coherent way, and where I have and exercise a capacity to have non-instrumental desires that fit together in a coherent way. Or I may have a non-instrumental desire that p because I believe something that that non-instrumental desire coheres with

4. Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1994), Chapter 5.

especially well—that I would non-instrumentally desire that p if I had a maximally informed and coherent and unified desire set, perhaps—where I have and exercise the capacity to have non-instrumental desires that fit together with my beliefs in a coherent way. These claims are, of course, controversial.<sup>5</sup> But the crucial point, for present purposes, is that we shouldn't so set things up as to preclude the truth of such claims. I will return to this at the end.

## II

*Instrumental desires are not distinct from the non-instrumental desires and means-end beliefs that explain them, but are rather just the complex state of having such non-instrumental desires and means-end beliefs standing in a suitable relation.*

Talk of the difference between instrumental and non-instrumental desires being a matter of what explains them might suggest, again misleadingly, that when non-instrumental desires and means-end beliefs give rise to instrumental desires, these instrumental desires are new desires with their own independent existence. But this is not the right way to think about instrumental desires.

It is a striking fact that instrumental desires disappear immediately an agent loses either the relevant non-instrumental desire or means-end belief. This is, if you like, part of what it is to be an instrumental desire, as opposed to a non-instrumental desire. Yet there is no reason why this should be so if an instrumental desire were merely a desire that has a non-instrumental desire and a means-end belief somewhere in its causal history. Why should a desire disappear when (say) the desire that caused it, way back when, disappears? Instrumental desires are thus better thought of as being nothing over and above the non-instrumental desires and means-end beliefs that explain them. When we said earlier, following Hume, that instrumental desires are those that can be explained by non-

5. Compare Geoffrey Sayre-McCord, 'The Meta-Ethical Problem: a discussion of Michael Smith's *The Moral Problem*' in *Ethics* 108 (1997), 55–83, and Michael Smith, 'In Defence of *The Moral Problem*: A Reply to Brink, Copp and Sayre-McCord' in *Ethics* 108 (1997), 84–119.

instrumental desires and means-end beliefs, what we had in mind was this constitutive claim, not a claim about causal history.

But nor should we think of an instrumental desire as just the mereological sum of a non-instrumental desire and means-end belief. An agent may have a non-instrumental desire and a means-end belief and yet have no corresponding instrumental desire. Someone could (say) have a non-instrumental desire to be healthy and believe that he can be healthy by exercising and yet have no instrumental desire whatsoever to exercise. Non-instrumental desires and means-end beliefs must be put together, as it were, if an agent is to have an instrumental desire. We should therefore suppose that an instrumental desire is nothing over and above a *suitably related* non-instrumental desire and means-end belief.

Against this, it might be thought that the systematic employment of the criteria we use when we attribute non-instrumental desires and means-end beliefs will require us to posit appropriate instrumental desires whenever people have relevant non-instrumental desires and means-end beliefs. When someone desires to be healthy and believes that she can be healthy by exercising, it might be thought that she must be ascribed a desire to exercise: that desiring to be healthy just is desiring to exercise when you believe that exercise causes health. But this is evidently untrue.

Desires and beliefs are multi-track dispositions—there are dispositions to act, dispositions to feel pleased or disappointed, dispositions to engage in various sorts of thought processes, dispositions to say certain things in response to questions, dispositions to use certain considerations in further reasoning, and so on and so forth—and the criteria that we use in the attribution of desires and beliefs are a weighted sum of these different factors. When we attribute non-instrumental desires and means-end beliefs to people there is therefore ample room for the hypothesis that these states are present, but simply not related to each other in the way required for the people in question have the relevant instrumental desires.

### III

*Agents who have non-instrumental desires and means-end beliefs have a more locally coherent psychology when they have*

*corresponding instrumental desires, and they have a more globally coherent psychology when the relative strengths of their instrumental desires covary with the strengths of their non-instrumental desires and the levels of confidence they have in the things about which they have means-end beliefs.*

Though, as we have just seen, there is no difficulty in attributing to someone a non-instrumental desire and a means-end belief but no relevant instrumental desire—the desire to be healthy and the belief that exercise causes health (say), but no instrumental desire to exercise—it must be said that such an attribution would be puzzling. The question is what this shows.

To desire something is to desire that some state of the world obtains. But it is a mundane fact that states of the world obtain as a result of the world's being some way or other: things happen as a consequence of other things. So anyone cognisant of this fact is in a position to see that to desire that a certain state of the world obtains is to desire that it obtains as a result of the world's being one of the ways it can be which will lead to the occurrence of that state of affairs. But now consider someone who desires that a certain state of the world obtains as a result of the world's being one of the ways it can be which will lead to the occurrence of that state of affairs, and who believes that the way the world can be which will lead to the occurrence of the desired state of affairs is the *p* way. For this person to fail to desire that the world be the *p* way is, it seems to me, quite literally for them to be in an incoherent state of mind. It is for them to fail to put the original desire and belief together in the way in which they are committed to putting them together. For, given their belief, their original desire is already, so to speak, targeted on the world's being the *p* way.

But even if this explanation of the normativity of the suitable-relatedness of a non-instrumental desire and means-end belief in terms of coherence is accepted, it might be thought that there is still room for disagreement about how extensive such requirements of coherence are. For example, suppose I have a non-instrumental desire to be healthy and that I believe there are two ways the world could be which lead to this result. I believe that it would result from exercise, or from my having more income, but I do not believe that I could exercise and increase my income at the same time. What does coherence require of me in this case?

After all, the world doesn't have to be the I-exercise way in order for it to be the I-am-healthy way, and nor does it have to be the I-earn-more way.

The answer, I think, is that coherence requires me to put my non-instrumental desire to be healthy together with each of these beliefs. For, given my beliefs, my non-instrumental desire is already targeted, so to speak, on each of these ways the world could be. I desire the realization of the possibility that I am healthy, and I believe that this possibility partitions into two sub-possibilities: the possibility that I exercise and the possibility that I increase my income. Following through the logic of the previous argument would thus seem to imply that coherence requires me to have both an instrumental desire to exercise and an instrumental desire to increase my income. Putting at least one of my means-end beliefs together with my non-instrumental desire will allow me to achieve a certain amount of local coherence, but I will achieve more if I put my non-instrumental desire together with both.

Moreover, sticking with this case, it seems that coherence also makes demands on the strengths of these instrumental desires. If, for example, I am equally confident about the two causal claims then coherence requires me to be indifferent between the two options: my instrumental desires will have to be equally strong. But if I am more confident of one than the other then it seems that, in order to satisfy the demands of coherence, my instrumental desire for the one about which I am more confident will have to be stronger. The effect of decreased confidence is, if you like, to dilute desire for that option. The upshot is thus that, even though I might meet coherence's demand that my non-instrumental desire be suitably related to my two means-end beliefs, I might still fail to meet coherence's further demand on the strengths of my two instrumental desires.

Coherence seems to make other more global demands as well. Suppose this time I have two desires, a non-instrumental desire to be healthy and a non-instrumental desire for knowledge, and that I believe all of the following: that exercise causes health, that reading causes knowledge, and that I cannot exercise and read at the same time. Finally, just to keep things simple, suppose I am equally confident about each of these things and that I have no further desires or beliefs. Does coherence require that both of my

non-instrumental desires stand in suitable relations to the relevant means-end beliefs, and hence that I have both an instrumental desire to exercise and an instrumental desire to read?

The considerations adduced above would seem to apply equally to the two non-instrumental desires. Coherence demands that my two non-instrumental desires be suitably related to each of my means-end beliefs: I would have to have both an instrumental desire to exercise and an instrumental desire to read. Moreover it once again seems that, though I might satisfy that demand, I might fail to meet a further demand that coherence makes on the strengths of these instrumental desires. If my non-instrumental desires for health and knowledge are equally strong then, in order to satisfy the demands of coherence, it seems that I will have to be indifferent between the two options: my instrumental desires to exercise and to read will have to be equally strong. But if one of my non-instrumental desires is stronger than the other then it seems that, in order to satisfy the more global demands of coherence, my instrumental desire for the one which leads to the outcome that I more strongly desire will have to be stronger.

Moreover there are also mixed cases. Suppose I have a stronger non-instrumental desire to be healthy and a weaker non-instrumental desire for knowledge, and that I believe that exercise causes health, that reading causes knowledge, and that I cannot exercise and read at the same time, but that I am more confident of the connection between reading and knowledge than I am about the connection between exercise and health. In that case it seems that, once again, coherence will demand that I have instrumental desires both to exercise and to read, but it also seems that coherence will require that the strengths of these instrumental desires depend on the strengths of the two non-instrumental desires and the levels of confidence associated with my two means-end beliefs. If my confidence is greater enough, then coherence may even require that the instrumental desire to read is stronger than the instrumental desire to exercise, notwithstanding the fact that the non-instrumental desire for knowledge that partially constitutes it is weaker than the non-instrumental desire for health which partially constitutes the instrumental desire to exercise.



## IV

*Actions are the product of instrumental desires that are complexes of suitably related non-instrumental desires and means-end beliefs of a particular kind.*

I have already alluded to the standard story of action explanation that we have inherited from David Hume.<sup>6</sup> According to this story, actions are caused and rationalised by a pair of mental states: a desire for some end, where ends can be thought of as ways the world could be, and a belief of the agent that something she can just do, namely move her body in a certain way, has some suitable chance of making the world the relevant way.

When we put this story together with the account we have just given of the nature of instrumental desires we get the following result. The pair of mental states that Hume says explain actions are none other than the non-instrumental desires and means-end beliefs that constitute instrumental desires when they are suitably related. Actions are therefore, according to the standard Humean story, the product of a *sub-class* of instrumental desires. They are the product of that sub-class where the means-end beliefs are about upshots of things that the agent can just do. Since this conclusion will be crucially important for what follows, let me say a little in defence of it.

Consider what must surely be the worst case for the claim that actions are the product of instrumental desires. Suppose an agent has a non-instrumental desire to do one of the things she can just do. Suppose, in other words, that she desires to perform what Danto calls a 'basic action', perhaps the basic action of moving her arm.<sup>7</sup> In that case it might be thought that no belief about what her moving her arm leads to could possibly be required to explain her action. No belief could be required because the agent is indifferent to the consequences. Her non-instrumental desire to perform the basic action of moving her arm must therefore suffice all by itself. Or so it might be thought.

6. David Hume, *A Treatise of Human Nature*, 1740 edition, (Oxford: Clarendon Press, 1968), especially Book II, Section III; see too Donald Davidson, 'Actions, Reasons and Causes' reprinted in his *Essays on Actions and Events* (Oxford: Oxford University Press, 1980).

7. Arthur C. Danto, 'What We Can Do', *Journal of Philosophy* 60 (1963), 434-445.

But in fact it seems that a means-end belief is still required, even in this case. For the agent still needs to believe that moving her arm will result *from something that she can just do*. In other words, she must put her non-instrumental desire together with her belief that she can perform the basic action of moving her arm and that doing so will lead to the desired result, namely, the movement of her arm. In order to see that such a means-end belief is required even in this case we must consider what happens when the non-instrumental desire to perform such a basic action is put together with the various possible combinations of belief and ability, as laid out in the following matrix.

		able to perform the basic action of moving my arm	
		yes	no
believe myself able to perform the basic action of moving my arm	yes	1	2
	no	4	3

(1) is the normal case in which I truly believe myself able to perform the basic action of moving my arm, and (2) is exactly the same, epistemically speaking, except that my belief is false: perhaps I have become paralysed without realizing it. But since, in both of these cases, if I happen to have a suitably strong non-instrumental desire to move my arm then I will presumably at least try to do so—the only difference will be that in (1) I succeed, whereas in (2) I fail<sup>8</sup>—neither is relevant to deciding whether a means-end belief is necessary for action. For in both cases I act, and in both cases I have the relevant means-end belief. Cases (3) and (4) are thus the real tests.

(3) is the case in which things are not as in the normal case and I am cognisant of this fact. Perhaps, being paralysed, I am unable to move my arm, and, being fully aware of my paralysis, I lack the belief that I am able to do so. This sort of case brings out the insight in the following remark of Anscombe's.<sup>9</sup>

8. Jennifer Hornsby, *Actions* (London: Routledge and Kegan Paul, 1980).

9. G. E. M. Anscombe, *Intention* (Oxford: Basil Blackwell, 1957), 52.

People sometimes say that one can get one's arm to move by an act of will but not a matchbox; but if they mean 'Will a matchbox to move and it won't', the answer is 'If I will my arm to move in that way, it won't', and if they mean 'I can move my arm but not the matchbox' the answer is that I can move the matchbox—nothing easier.

For even if, in case (3), I have a suitably strong non-instrumental desire to perform the basic action of moving my arm, my arm will be, for all intents and purposes, exactly like a matchbox, and my non-instrumental desire will therefore be, for all intents and purposes, just a like a non-instrumental desire I might now conceive to perform the basic act of moving a matchbox. Just as I am at a loss to know what it would mean even to try to move the matchbox in that way, I will be at a loss to know what it would mean even to try to move my arm in that way. Moreover the explanation would seem to be the same in each case, namely, that I lack the crucial means-end beliefs: that is, I lack the belief that there is something I can just do will that lead to the desired outcome. (3) is thus no counterexample to the necessity of the relevant means-end belief.

Having said that, what should we make of the case in which, being unable to (say) wiggle my ears, and believing myself unable to do so, I am none the less invited to try? 'Just do this!' I might be told by someone who demonstrates. I take it that in this sort of situation we do ordinarily succeed in at least trying to wiggle our ears. But once it is conceded that such an attempt is possible then it might be objected that, given that I don't have the belief that I can perform the basic action of wiggling my ears, that just proves that action is possible when I lack the relevant means-end belief.

But though it is true that I can try to wiggle my ears in this sort of situation I crucially do not try to do so by acting on my non-instrumental desire to perform the basic action of wiggling my ears. Rather what happens is that I see that the person who is wiggling their ears is doing something that I can just do—raising their eyebrows (say)—and what I try to do is to wiggle my ears by means of doing that. In other words, my action of trying to wiggle my ears results from my non-instrumental desire to perform the non-basic action of wiggling my ears and the belief I have as a result of the demonstration I've been given that something I can just do, namely raise my eyebrows, will lead to

this result. So this case is no challenge to the necessity of the means-end belief either. It is, again, a case in which I have the relevant means-end belief.

The situation just discussed helps us understand what is going on in case (4). In (4) I have an ability to perform a certain sort of basic act, but I don't believe that I have this ability. Is it possible for me to act on the non-instrumental desire to perform the basic action alone in this sort of situation? The answer is, once again, that it is not. Lacking the belief that this is something I can just do, it seems that I would once again be at a total loss to know what it would mean even to try to just do the thing in question. In other words, since case (4) is epistemically just like case (3), case (4) presents no challenge to the necessity of the means-end belief either.

Once again, however, we should immediately acknowledge that it is possible for me to come to realise that I have the ability to perform a certain sort of basic action that I hadn't hitherto believed myself able to perform. Consider again the case in which someone asks me to wiggle my ears and then gives me a demonstration of the way in which they do it, but imagine this time that, unbeknown to me, I am one of those who is able to perform the basic action of wiggling my ears. It is certainly possible that I would succeed in trying to perform the non-basic action of wiggling my ears by copying what they do—by raising my eyebrows (say)—and it is also possible that, through subsequent experimentation, I could come to discover that I can wiggle my ears as a basic action, that is, without doing so by means of raising my eyebrows. But if this is right then this case presents no counter-example to the claim that a means-end belief is necessary for action either. It simply underscores the fact that our beliefs about which basic actions we can perform are derivable from experience.

## V

*Agents meet the standards of instrumental rationality to the extent that they satisfy the requirements of local and global coherence on their non-instrumental desires and means-end beliefs.*

We are now in a position to say what it means for agents to meet the standards of instrumental rationality. My suggestion,

perhaps unsurprisingly, is that agents meet such standards when they have non-instrumental desires and means-end beliefs that satisfy the requirements of local and global coherence as described above. In other words, their non-instrumental desires and means-end beliefs must be suitably related to each other—they must have instrumental desires corresponding to each of their non-instrumental desire/means-end belief pairs—and the relative strengths of their instrumental desires must be the right kind of product of the strengths of their non-instrumental desires and the confidence with which they hold their means-end beliefs. If this is agreed then a number of important consequences follow.

To begin, note that our defence of the normativity of instrumental rationality—the defence of the claim that having suitably related non-instrumental desires and means-end beliefs is required for psychological coherence—didn't presuppose any deep nexus between instrumental rationality and instrumental reasoning. Nor should this be surprising. For whereas reasoning is a process that takes an agent from one set of beliefs to another set of beliefs, the transitions with which we have been concerned take an agent from non-instrumental desires and a means-end beliefs to instrumental desires. There has therefore been no suggestion that an agent who has a non-instrumental desire and a means-end belief, and who goes on to form a corresponding instrumental desire, must do so by engaging in a process of reasoning. To desire something simply is not to accept a proposition that is fit for use in subsequent reasoning. Reasoning thus doesn't seem to be relevant to the task.

Of course, all sorts of reasoning processes may, as it happens, be employed when I succeed in being instrumentally rational. If (say) I have a non-instrumental desire to be healthy and I also have sufficient self-knowledge to realise this, then, if I also believe that exercise causes health, then I will be in a position to reason myself to the conclusion that exercising will cause me to get something that I non-instrumentally desire, namely health. And, having formed this belief, I will be in a position to reason myself to the further conclusion that the formation of an instrumental desire to exercise is required for local psychological coherence. Finally, equipped with this belief, I might find myself forming an instrumental desire to exercise. But it is important to emphasise the contingency of the connection between the reasoning process

described and the formation of the instrumental desire. Forming the beliefs described is neither necessary nor sufficient for forming the instrumental desire to exercise.

A second important consequence concerns the scope of instrumental rationality. As we saw above, though all actions are the product of instrumental desires, these instrumental desires are a proper sub-class of the full set of such desires. Instrumental desires are simply suitably related pairs of non-instrumental desires and means-end beliefs, but actions are the product of that sub-class of instrumental desires that are constituted by suitably related pairs of non-instrumental desires and means-end beliefs to the effect that something that the agent can just do will have the desired upshot. It thus follows that an agent could be instrumentally irrational for failing to have an instrumental desire, but that the instrumental desire in question has nothing whatsoever to do with the desired result's happening as a result of anything that she can just do. Since such instrumental desires will not feed in any direct way into the agent's choices or intentions, it follows that the scope of instrumental rationality extends way beyond choice and intention.

For example, suppose I desire that my great-great-grandfather didn't suffer as a result of hearing about a crime committed by my great-grandfather, and that I have two relevant instrumental beliefs, that is, two beliefs about how this could have come about. I believe that he won't have suffered if he died before my great-grandfather's crime was committed, and I believe that he won't have suffered if, though he was alive, people tried to keep the facts from him. If the requirements of instrumental rationality are just the coherence requirements described earlier then it follows that I am instrumentally irrational if I fail to apportion my desire to these two ways in which I believe that the state of the world that I non-instrumentally desire could come about. For how could I coherently desire that my great-great-grandfather didn't suffer as a result of hearing about a crime committed by my great-grandfather, and yet be indifferent to the world's being the ways I believe it would have to be for that to be the case? Moreover, even if I do have both of these instrumental desires, the account we have given of the global coherence constraints on non-instrumental desires and means-end beliefs entails that I am instrumentally irrational if, having more confidence that my

great-great-grandfather will not have suffered if he died before the crime was even committed than if people tried to keep the facts from him, my instrumental desire that the former is the case is not stronger than my instrumental desire that the latter is the case. But, and here is the crucial point, none of these instrumental desires constitutes a choice or an intention. Indeed, since the desires in question all concern the past they don't seem to have anything very obvious to do with my choices or intentions.

A third important consequence concerns the evaluation of agents. As we saw above, all actions, even those whose only purpose is to perform some basic action, are the product of a suitably related non-instrumental desire and means-end belief about the upshots of things that the agents can just do, where being suitably related is a requirement of local coherence on those non-instrumental desires and means-end beliefs. It therefore follows that anyone who acts must, thereby, meet that minimal standard of local coherence: anyone who acts is instrumentally rational.

It might be thought that this explains the appeal of the following remark of Davidson's.<sup>10</sup>

In the light of a primary reason . . . [where a primary reason is just a suitably related desire and means-end belief] . . . an action is revealed as coherent with certain traits, long- or short-termed, characteristic or not, of the agent, and the agent is shown in his role of Rational Animal.

But the rhetoric turns out to be rather overblown. It is, after all, consistent with the local coherence required for action that an agent's non-instrumental desires and means-end beliefs fail to meet more global requirements of coherence. For example, given the strengths of their various non-instrumental desires, and the confidence with which they hold their means-end beliefs, agents might not have the instrumental desires required for global coherence. Certain instrumental desires might be missing—in other words, certain of their non-instrumental desires and means-end beliefs might not be suitably related—and, even if not missing, certain of their instrumental desires might be weaker or stronger than is required for global coherence.

10. Davidson, 'Actions, Reasons and Causes', 8.

The upshot is that it is only with a rather large pinch of salt that we can accept Davidson's suggestion that when someone acts, he is thereby shown in his role of Rational Animal. A more truly rational animal will not just meet the standard of local coherence required for action, but will have a full complement of suitably related non-instrumental desires and means-end beliefs, where the strength of the resultant instrumental desires is that required for global coherence. Such an animal will be more truly rational because it will not only be capable of action, but will be in a state of maximal preparedness to act in ways that optimally satisfy their desires, given their beliefs, under a whole range of counterfactual circumstances in which their non-instrumental desires and means-end beliefs are different.

A fourth and final consequence concerns a limitation on the ambitions of an account of instrumental rationality. The explanation of the normativity of instrumental rationality offered here makes it plain why agents could satisfy all of the requirements of instrumental rationality and yet still be vulnerable to rational criticism. For the demands of both local and global coherence described take as given the fact that agents have the non-instrumental desires that they have, they take as given the fact that these desires have the strengths they have, they take as given the fact that agents have the means-end beliefs they have, and they take as given the fact that their means-end beliefs are held with the confidence with which they are held. But in that case it is consistent with an agent's satisfying all of the demands of instrumental rationality, as these have been conceived of here, that their psychology fails to satisfy even more global requirements of coherence on the elements that are taken as given.

What might these even more global requirements of coherence be like? Certain of an agent's means-end beliefs might fail to cohere with their evidential beliefs, beliefs that provide them with ample evidence that their means-end beliefs are false. Or their non-instrumental desires might fail to cohere with each other. Or certain non-instrumental desires might fail to cohere with their beliefs about (say) which non-instrumental desires they would have if they had a maximally informed and coherent and unified desire set. In these ways, and perhaps in others too, the non-instrumental desires and means-end beliefs agents have might themselves become vulnerable to rational criticism.



The point is not, of course, that the account of instrumental rationality offered here entails that an agent's non-instrumental desires and means-end beliefs are vulnerable to such criticism. The point is rather that, at least as the requirements of instrumental rationality have been conceived of here, such rational criticism isn't ruled out. Whether or not such rational criticism is possible turns, in the end, on whether there really are the even more global requirements of coherence, like those suggested, that would underwrite the possibility of such criticism. Moreover this is, it seems to me, just as it should be. Though I do not myself believe that instrumental rationality is all there is to practical rationality, an account of the normativity of instrumental rationality should make vivid the temptation to suppose that it is.