



THE MYTH OF INSTRUMENTAL RATIONALITY

BY JOSEPH RAZ

JOURNAL OF ETHICS & SOCIAL PHILOSOPHY

VOL. 1, NO. 1 | APRIL 2005

URL: WWW.JESP.ORG

COPYRIGHT © JOSEPH RAZ 2005

The Myth of Instrumental Rationality

Joseph Raz¹

Recent interest in the nature and presuppositions of instrumental rationality was inspired to a considerable degree by arguments designed to show that it presupposes other forms or kinds of rationality, or (to put it in the nonequivalent way in which the point is more commonly put) that claims that there are reasons to pursue the means to our ends presuppose that the ends themselves are worth pursuing, or that there are adequate reasons for pursuing them. The discussion of instrumental rationality is bound up with discussions of instrumental reasons and of instrumental reasoning that guides deliberation, and which, other things being equal, it is irrational knowingly to flout. The interest in understanding instrumental rationality was thus at least partly a result of an interest in – often hostility to – an ambitious claim, namely that all practical reasons are instrumental, that practical normativity is about the normativity of following the means to our ends.

I say little about this issue. My main aim is to explain the normative character of the phenomena that are commonly discussed when theoretical writers discuss instrumental rationality and instrumental reasons. The discussion will assume that there are forms of practical normativity, of practical reasons, which are not instrumental in nature. The question central to the inquiry is what, if any, normative difference does adopting or having an end make? For example, are there instrumental reasons and, if there are, how do they relate to having ends? Are instrumental reasons distinctive kinds of reasons, whose normativity differs in its underlying rationale from that of, say, moral reasons, or of other kinds of reasons? Similarly, is instrumental rationality a distinct form of rationality?

Reflecting on these questions, we are liable to be torn both ways. On the one hand, we feel that the value of the means derives from the value of the ends. If there are reasons to take the means, they must be none other than the reasons to pursue the ends, or at least they must derive from them. On the other hand, we also feel that failure to take the means to one's ends is a distinct kind of failure, different from the failure to have proper ends, or to value them properly. The first response – that there are no distinctive instrumental reasons – is reinforced by the thought that, say, a would-be

¹ I am grateful to John Brunero, Ulrike Heuer, Philip Kitcher, Scott Shapiro, Mike Martin, Chris Peacocke, Seana Shiffrin and Mathew Smith for comments on parts or complete drafts of this paper. The paper was presented to the Warren Quinn Conference at UCLA in November 2004, where I had the benefit of comments by Tamar Shapiro.

murderer cannot create for himself a reason for poisoning his intended victim just by making it his goal to kill him. The second response, however, is reinforced by the thought that a person with evil or worthless ends who fails to take the proper means to his ends is, perhaps luckily, irrational in a way that is indifferent to the character of his ends. He is irrational in the same way as someone whose ends are worthy, but who fails to take the means toward them.

In the first section, I will suggest that (subject to some qualifications) we have (instrumental) reasons to facilitate the realization of anything of value whose realization is not deeply impossible, regardless of whether or not its realization is one of our ends. In section 2, I agree with those who (a) deny that our ends (and intentions) are reasons to take the means for their realization and (b) argue that, nevertheless, our ends affect what is rational for us to do. I do, however, contend that some attempts to square this circle failed, and offer a different solution. Section 3 defends the proposed position by arguing that we do not have reason to avoid contradictions as such. Section 4 complements the earlier discussion by pointing to additional practical implications of ends.

The conclusions of these sections, while following in the footsteps of a number of recent writers on the subject, do deviate from long-established ways of treating the phenomena philosophers discuss under titles such as instrumental reasons, normativity or rationality. Section 5 faces the oddity of my conclusions. It does so by challenging the thought, common to some writers who otherwise vary considerably, that instrumental rationality or instrumental reasons or both are a distinctive type of rationality or of reasons. In particular it reinforces the critique, implied in the preceding sections, of the view that instrumental reasons or instrumental rationality are a basic, perhaps simpler, domain of reasons or rationality that, according to some, exhausts practical reasons and practical rationality, whereas, according to others, it constitutes the most elementary and theoretically least problematic part of.

1. Instrumental reasons without ends

I start with informally introducing some of the terminology I will use and the theses I will presuppose. Imagine some good, valuable, worthwhile objects, states or activities, whose value is not (or not exclusively) due to the value of their actual or probable consequences, or consequences their existence facilitates (e.g., by being pre-conditions for their existence, or tools that can be used to bring them about). Say a beautiful musical composition, canoeing down the Colorado River, the state of understanding Kant's philosophy in all its details. I will proceed on the assumption that (some exceptions that need not concern us aside) their value is a reason (which we

call “non-instrumental”) for everyone to engage with and to respect what is of value, provided only that doing so is not deeply impossible.

There is no quick explanation of what it is to engage with something of value, or what constitutes respecting it. It all depends on what kind of value it has. If it is a concert, engaging with it could be to listen to it, or play in it, or conduct it. If it is a dance or hill walking, it may be joining in; if it is knowledge of Kant, then it may be a matter of reflecting on Kant’s views, comparing them with others, raising issues and responding to them – in short making one’s knowledge active, making it play a part in one’s life or work. Respect for what is valuable is expressed by avoiding harming it, protecting it from harm, and so on. For present purposes, we can rely on examples such as these to guide us.²

It is not the case that I have reason to attend the symposium that was recorded by Plato in the dialogue of that name, nor is it the case that I have reason to travel outside the solar system, but I have reason to go to the concert conducted by Abbado tonight even though I cannot, for it is sold out. I do not know how to distinguish in the abstract between reason-negating and other impossibilities. Perhaps it is just a matter of the likelihood that the impossibility will lift, or that it is temporary.

The reasons that will be discussed here are, of course, normative reasons³, and they are what we often call *prima facie*, or “other things being equal” reasons. They can be defeated by other reasons in a variety of ways. Even when those reasons are not defeated, the agent may have no worse reasons to engage in other activities or bring about other states or create other objects. Only when there is only one option, the reasons for which defeat all competing reasons, do we have conclusive reason to take that option, and the reasons for the option, taken together (or any combination of some of them that is sufficient to defeat all conflicting reasons), constitute a conclusive reason for it.

Some people do not share the assumption that the (non-instrumental) value of things is a reason to engage with them and to respect them. So long as they have other ways of explaining how it is that people may have reasons to act that do not derive exclusively from reasons they have to bring about the consequences of their actions, this disagreement may not matter to the arguments to come.

² See more generally my *Value, Respect and Attachment* (Cambridge U.P. 2000), Ch. 4.

³ “Reasons” has, of course, a second meaning, designating, roughly speaking, the facts that explain whatever they are reasons for (as in “the reason for the storm was ...”). When what is explained are phenomena that can be shaped by normative reasons, those, and belief in them, are often invoked as explanatory reasons – a fact that does not conflict with the distinctness of meaning and of reference of normative and explanatory reasons.

I am assuming that rationality has to do with the perception of what are or are taken to be normative reasons, and the response to them. To avoid tedium I will refer informally to normative and rational principles, considerations, etc., assuming that in the contexts, their inter-relations are easy to discern.

Some things that we have reason to do we do simply by doing them. In doing so we may also be doing something else. Davidson's familiar example will serve: We may turn on the light by flicking the light switch.⁴ I do not flick on the light switch by doing anything else.⁵ I just flick it on. I do turn on the light by flicking the light switch, the second act being, in the circumstances, a constituent of the first, of my turning on the light (though not of the action-type of turning on the light). Often, however, things are different, and we can perform an action only by performing another action first, one that is not a constituent part of the main action that we can then proceed to perform. I can travel to Oxford only if I first stop writing this paper, get out of my chair and put on my shoes. None of them are elements of going to Oxford, but they are, in the circumstances, preconditions for doing so. Other actions may be means for doing so. For example, filling my car with petrol (if I intend to drive) or buying a ticket (if I intend to go by train). One can spend a happy time distinguishing means from preconditions from facilitating conditions and so on. I will avoid that, and will refer to all of them as facilitating steps or acts.

Facilitating steps can come in ordered sequences, each constituting one way of bringing us to the point where we can take the action we have reason to take. We could call each a (possible) plan⁶ of how to get to a point at which we can take that action.⁷

When we have an undefeated reason to take an action, we have reason to perform any one (but only one) of the possible (for us) alternative plans that facilitate its

⁴ Davidson, "Actions, Reasons and Causes" (1963) reprinted in *Actions and Events* (Oxford U.P. 1980), p. 14. Needless to say, my use of it is not Davidson's.

⁵ Some would insist that only actions consisting in nothing more than bodily movements caused by the appropriate intention can be acts not done by doing others. It seems to me, however, that that is an artificial stipulation. The fact that normally we do not intend to move our fingers, but rather intend to flick the switch (as well as, of course, to turn on the light) is a good ground for taking flicking the switch not to be an action done by performing another action, even though I do it by moving my fingers (which was intentional because I intended to flick on the switch in that way).

⁶ Needless to say plans can overlap: If the nearest bus stop to Oxford is at the train station, then going to that location is part of both the plan to get to Oxford by bus and the plan to get there by train.

⁷ On various senses of "plans," see Bratman, *Intention, Plans and Practical Reason* (Harvard U.P. 1987, chapter 3). In this paper "a plan" refers to an abstract structure in the nature of a recipe.

performance. If I have an undefeated reason to be in Oxford, I have reason either to get out of the chair, fill my car with petrol and drive to Oxford, or to take the train or take the bus or beg a lift from a friend, etc. I will call this the *facilitative principle*, and the reasons it talks of – that is the reasons for taking facilitating actions – facilitative reasons. I will call the action whose desirability is the reason for taking facilitating actions the source action, and the reasons for it the source reasons. The desirability of the source action is a reason to adopt a plan that will facilitate it. The existence and force of the source and of the facilitative reasons do not depend on their existence or force being recognized by the people for whom they are reasons.

It seems plausible to assume that much of what people have in mind when they talk of *instrumental reasons* is captured by this principle. Perhaps, and I will come back to this question, there are other so-called instrumental reasons. But it seems plausible that, if the principle is sound, then the reasons it singles out are instrumental reasons, as this term is commonly used. If this is so then we have instrumental reasons independently of having any ends. The facilitative principle is indifferent to what are our ends and asserts that we have instrumental reasons to facilitate any worthy state or action, even those whose performance is not among our ends, even those that we cannot pursue if we pursue our ends.

The reasons we have for facilitative acts show their instrumental nature by their dependence on the source action, which makes them more dependent on the circumstances of action. No occasion of possible realization is presupposed by true propositions of (non-instrumental) reasons we have. I have a reason to see the film *The Seventh Seal* even though I do not know when, if ever, it will be possible for me to do so. The statement is true even if conformity is not possible, except, as was remarked above, that it excludes deep impossibility. Statements of instrumental reasons, that is of reasons to take facilitative actions, presuppose an occasion (or class of occasions) on which the facilitative actions will be possible, and will facilitate: If they could facilitate under different conditions but would not facilitate today, then I do not have reason to take them today. We need to distinguish between plans and actions within a plan. We have reason to adopt and pursue a plan only if, and for as long as, it is feasible and affords reasonable chances of successfully facilitating realization of the action it is designed to facilitate. That a plan would have been effective under slightly different circumstances is neither here nor there.⁸ Steps within a plan we have only conditional reason to take, the condition being that we have adopted and are pursuing the plan (and that it is still reasonably likely to facilitate what it is meant to facilitate).

⁸ Though, of course, there can be an alternative plan that consists of this plan, preceded by steps to make it realizable.

So, for example, if I have reason to visit my grandmother who lives on Easter Island today, I also have a reason to buy a flight ticket to go there today, but only if she is there or will be there, and only if it is possible to get there by air today. If there is an air strike, I have no reason to pursue the plan of flying there today, and therefore no reason to buy a flight ticket. Such actions will not facilitate my visiting her,⁹ even though I still have a reason to visit her today, albeit one with which (let us assume) it is impossible for me to conform.

Another way in which facilitative reasons depend on circumstances is that we have facilitative reasons only in circumstances in which the source reason is not defeated. Imagine that I have to choose which of three films, A, B or C, to go to. I can see any, but no other option is available. Imagine further that all three are worth seeing. So I have reason to see each of them, but cannot see more than one today. Finally, let it be assumed that, for some reason, I should see A or B rather than C, but there is no better case for seeing A rather than B, nor the other way round. In this situation, I have reason to facilitate seeing A as well as reason to facilitate seeing B, but no reason to facilitate seeing C. For example, I have reason both to buy a ticket for A, and to buy a ticket for B, but no reason to buy a ticket for C.¹⁰ Because it is a good film I still have reason to see C. But because in today's circumstances that reason is defeated, I have no reason to facilitate action conforming to it. This is nothing to do with seeing C not being my end. I have no end, no decision what to do.

Suppose that I resolved my doubts and do intend to see film A. I have reason to facilitate doing this, but do I still also have reason to facilitate seeing film B? The answer depends on whether my intention means that now the reason to see A is stronger than the reason to see B. If it is, the reason to see B is defeated, and I do not have reason to facilitate seeing B. If, however, my intention does not affect the stringency of the reasons for either option, I still have reason to facilitate both. You may ask: since I cannot see both why do I have reason to facilitate both? Given my intention I will not do so, rather I will buy a ticket for A. That will affect my reasons about which film to see. I will have a better reason to see the film I have a ticket for. This means, according to the facilitative principle, that I no longer have a reason

⁹ This appears to be a special case of the following: If it is certain that I will not do the source action, whether or not I can do it, there is no reason to take the facilitative action.

¹⁰ It may be said that I have reason to do that, too, as insurance in case I am wrong in thinking that I can see either A or B. Without denying this, I will continue in this paper to ignore reasons arising out of epistemic considerations, of which this is an example. Also, note that because I can only see one film, I have no reason to buy tickets for both A and B. The reasons are dependent and conforming with one cancels the other.

to facilitate seeing B, as the reason to see B has been defeated. If, however, I buy a ticket for B, I have reason to change my intention, as now the case for seeing B is stronger than the case for seeing A.

The stringency of the (overall) reasons for pursuing different plans facilitating one source action is not the same, and is not the same as the stringency of the reason for taking the source action. Since there may be diverse disadvantages and diverse additional advantages to pursuing different plans, the reasons for the different plans differ, and because they may have disadvantages that the source act does not suffer from, the stringency of the reasons for them may differ from the stringency of the reason for the source act. Furthermore, the reason to pursue any of the plans does not entail an unconditional reason to take any of the steps that are parts of it. But there is a conditional reason to take such steps under certain circumstances. The condition is not easy to specify and I will not try to do so now.

The hard question is: Is the facilitative principle sound? The simplest argument is that it is good to facilitate worthwhile actions, and I have been assuming all along that, at least if we can, we have reason to do what is good. This argument may appear too quick and question-begging. Another way to approach the issue is to imagine a situation where something clearly should be one's end, though it is not. Suppose Annabel has, in her care, a 2-year-old child, her child, and she should have been feeding it properly as one of her ends. As it happens, Annabel cares for nothing but her career, and feeding her child means nothing to her (though she will feed it if its cries disturb her peace, or to avoid being prosecuted for child abuse or neglect). Thinking of such a case, it seems difficult to deny that she has reason not only to feed the child but to buy food, which will facilitate feeding him. It seems to make no sense to think that she has a reason to feed the child, but no reason to make it possible for her to do so. That feeding the child is not an end of hers seems to be neither here nor there.

But if this is so regarding what should be her end, why is it not also true of valuable pursuits that she has reason to make her ends, but will not be wrong not to do so? Suppose that reading Conrad's *Lord Jim* is such a pursuit, but she could instead spend her free time reading Hardy's *Far From the Madding Crowd*, or watching the films of Orson Welles. That she may read Hardy rather than Conrad does not mean that she has no reason to read Conrad. Of course she does. If so, does she not also have the reason to take steps to make it possible to do so? As explained, if she pursues a permissible alternative course of action, there may be nothing wrong in her not taking such steps. If she intends to read Hardy and borrows his novel, she need not (for now) also borrow Conrad's. But until and unless her intentions and her pursuit of means change her reasons, she has reason to facilitate whatever ends she has undefeated reason to adopt.

Yet another approach to the facilitative principle is to think of the nature of reasons. One aspect of reasons for action is that they make choices,

intentions and actions intelligible both to the agent and to others. If Betty were to decide to live in L rather than B on the ground that L, being a metropolitan city, has many more cultural activities than B, her choice would be intelligible. The fact that modern dance is the only cultural activity she regularly takes an interest in (the only cultural pursuit that in some form is among her ends) is neither here nor there. It is intelligible that she would want to choose her town of residence in a way that facilitates pursuit of a wide range of cultural activities. If so, then in making such a choice she would be acting for a reason, and that reason is of course one that is available to all and would make any person's choice intelligible were they to make such a choice. Hence they all have the facilitative reason whether or not they decide to follow it, and whether or not they should conform to it.¹¹

2. The difference ends make

I have argued that at least some instrumental reasons do not presuppose having ends, that we have instrumental reasons that are unrelated to ends which we do have. It does not follow that our ends do not make a difference to the reasons we have. They can constitute or generate additional reasons, or they can affect the character or stringency of reasons we have anyway. If they do either they may similarly affect instrumental reasons that we have.

What are people's ends? For a start, they are not what they want or desire. There are many things that people want, but achieving or getting them is not among their ends. This may be because their desire is to achieve or be given something without trying, or because they do not think they stand a chance of getting it, or because they want other things more and think that they cannot get both, or they may be just too lazy, diffident or self-doubting to pursue it, to have it as their end, or they may believe that theirs are unworthy desires or that the objects of their desires are unworthy, or that they are under a duty not to pursue them. There are many other ways and factors that stop people from pursuing many of their desires, or from having

¹¹ Ulrike Heuer suggested to me that such cases are simply ones in which one has "reason to make sure that means are available if one were to adopt new ends." This seems to me close to the truth, and consistent with my contentions here. The point I am making is that such reasons do not presuppose a prior end that they facilitate, such as the end to find ends in the future. Note though that it is not so much that we have reason to facilitate adopting new ends, as reasons to facilitate acting for adequate reasons. The adoption of ends or forming of intentions is merely the way we come so to act. Thus formulated we can more clearly see that those reasons do not depend on having an end to act for adequate reasons in the future, including actions that we are not yet resolved to take. Rational agents, in being rational, are open to reasons that they have not considered before, or to which they did not yet respond.

their satisfaction or the realization of their objects among their goals or ends.¹²

Several recent writers agree that it is not the case that our having ends, whatever their content, is a reason for us to pursue them, or that they are or generate reasons to take facilitative steps toward their realization.¹³ I have already indicated the difficulty: People may have ends that, due to ignorance, weakness of the will, rashness or carelessness, or mistaken moral convictions, they should not have, ends that are worthless or even evil. Can they have reason to facilitate their realization just because they are their ends? It is no reply to this question that one's end is only a *prima facie* reason. That one's end is to kill some innocent person is no reason at all to buy poison or take any facilitating measures toward killing him. The view I defend here is of this family. I will not argue directly that our intentions or ends are not reasons. Rather, given the above objection to the view that they do in and of themselves provide reasons for action, I will try to show that the considerations that lead us to think that intentions and ends provide reasons can be, and should be, explained in ways without that assumption.

For example, the fact that people who pursue their ends act intelligibly, even when there are no adequate reasons for them to have the ends they pursue, does not require us to assume that their ends are their reasons, and that that makes their pursuit of them intelligible. To have an end

¹² A similar point is made by R.J. Wallace in "Normativity, Commitment, and Instrumental Reason" *Philosophers Imprint*, Vol. 1 No. 3 (2001), p. 14. None of this should be read to mean that people cannot have ends they believe to be unworthy, or contrary to their duties, etc. Contrast, e.g., J.L. Mackie: "If you want X, do Y" (or 'You ought to do Y') will be a hypothetical imperative if it is based on the supposed fact that Y is, in the circumstances, the only (or the best) available means to X The reason for doing Y lies in its causal connection with the desired end, X; the oughtness is contingent upon the desire." *Ethics* (Penguin Books 1977), pp. 27-28, or J. Dreier: "If you desire to A and believe that by phying you will A, then you ought to phye" in "Humean Doubts about the Practical Justification of Morality" in *Ethics And Practical Reason*, edited by Cullity & Gaut (Oxford U.P. 1997) p. 93.

¹³ See M. Bratman, *Intention, Plans and Practical Reason*; Broome, "Normative Requirements" *Ratio* 12 (1999), 398, Reprinted in *Normativity*, edited by J. Dancy (Blackwell, 2000), pp. 78-99, and "Practical Reasoning," in *Reason and Nature: Essays in the Theory of Rationality*, edited by José Bermúdez and Alan Millar, (Oxford U.P.); Korsgaard, "The Normativity of Instrumental Reason" in *Ethics And Practical Reason*, edited by Cullity & Gaut (Oxford U.P. 1997); R.J. Wallace, *ibid.* My agreement with this view does not contradict the view that I defended elsewhere, namely that there are personal commitments or resolutions that are reasons; indeed what I called protected reasons. They are the personal analogues of institutional decisions which, in a sound institution acting within its jurisdiction, are protected reasons for carrying out the decisions. See, *Practical Reason and Norms* (1975, current edition by Oxford U.P. 1999), pp. 65-72.

involves believing that it is worth having (at least other things being equal).¹⁴ That belief explains why people pursue ends. They take what they believe to be reasons for the ends as reasons to pursue the ends, and, as explained by the facilitative principle, to take steps to facilitate them.¹⁵

Even so, there appears to be a strong case for thinking that ends and intentions provide reasons. After all, people who fail to pursue the means to their ends display, or at least appear to display, a distinctive form of irrationality. “Have you not resolved to buy this car” (we are hardly likely to ask: Isn’t buying it an end of yours, but it comes to the same thing)? “Yes,” she says. “Do you realize that you cannot buy it without making an offer to buy it?” “Of course,” she says, “I am no fool.” “So, when will you make an offer?” “Oh,” she replies, “I have no intention of ever making an offer.” Something has gone wrong, and that is so whatever the merit or demerit of her buying the car. It appears that just by failing to intend to pursue the means to her end she is behaving irrationally.¹⁶ But if her end is no reason for her to facilitate its realization, how can that be?

One approach to the relevance of ends has been advanced and explored by J. Broome and by R.J. Wallace, who explains:

The principle [that he, following Korsgaard, calls the instrumental principle] imposes rational constraints on the attitudes of agents without entailing either that they have reason to take the means necessary relative to their ends, or that they are rationally required to believe that they should adopt the necessary means. (*ibid.*, 16)

Perhaps those who think that our ends or intentions create reasons for taking the means toward their realization believe that all rational constraints consist in reasons that constrain what we may rationally do or believe. That is their mistake.

But how does rationality constrain agents’ attitudes? By “governing combinations of attitudes” (*ibid.* 17). Broome has explored this idea in detail (and in summarizing his position, I adapt it to apply to the way the issues have been presented here). He points out that some conditional ought-propositions do not allow for detachment. Even if one ought (to do A if P) and P it does not follow that one ought to do A. To conclude that one ought to do A is to detach where detachment may not be warranted. For example, if P entails Q then, according to Broome, one ought (to believe that Q if one believes that P). But it does not follow that if one believes that P one ought to believe that Q. P may be false (or, if you think this makes a difference,

¹⁴ The qualification allows for having ends against one’s better judgment.

¹⁵ Though the reasons may change over time and may come to include the belief that one is trapped with an end one would be better without, but cannot give up because its pursuit incurred obligations to others, or for some other reason.

¹⁶ Ignoring this point seems to vitiate Korsgaard’s account (*ibid.*) of “the instrumental principle.”

obviously false), in which case the fact that it entails Q and that one believes that P is no reason at all to believe that Q.

Therefore, from

(1A) One ought (to do M if one intends to do E and M is a means to E), and

(2A) One intends to do E and M is a means to E, it does not follow that

(3A) One ought to do M

Following Broome, one may be tempted to claim, though he himself does not consider this matter, that the same is true regarding some conditional reasons. It does not follow from one has reason (to do A if P) and P that one has reason to do A. Similarly, it does not follow from:

(1B) one has reason (to do M if one intends to do E and M is the means to E); and

(2B) one intends to do E and M is the means to E; that

(3B) One has reason to do M

Adapting to Broome's reasoning, we can say that either (1A) or (1B) expresses the rational constraint imposed by the instrumental principle, and, since they do not entail (3A) or (3B), one's ends are not and do not generate reasons for action.

The problem is that (3B) is entailed by both sets of premises. Forget about ends. Reasons are reasons to do what will constitute conformity with the reason. (1B) does state that one has a reason. What is it a reason for? A roundabout way to identify the reason is to say that it is a reason to avoid being in a situation in which one would be in breach of that reason. And one would be in breach of it if one both intends E and fails to do M. There are two ways to avoid being in that situation. One is to abandon the intention to do E. The other is to do M. So one has both a reason to do M and a reason to abandon one's intention to do E (though no reason to do both, because once one does one of them the reason to do the other lapses). That means that, so long as M is the means to E and one intends to do E, one has reason to do M. By doing M, when it is the means to E, one acts in a way that puts one on the right side of reason. By doing M, one conforms to the reason stated in (1B). It follows that one has reason to do M.¹⁷

Can one argue that Broome is right and *the reason* (to take the means) is not one's intention, or one's having an end? This seems like hair splitting given that even if one does not have a reason to do E, once one intends to do it, one has a reason to M (which is the means to doing it). This suggests that it is not only natural, but also true and not misleading to say that the intention gives one a reason to pursue the means to its fulfilment.

¹⁷ My argument here is incompatible with accepting a rule of inference allowing introducing a disjunction within the scope of a modal operator (i.e. if "M" is a modal sentence-forming operator it licences inferring M(P or Q) from M(P). I believe this rule to be unintuitive and unjustified.

In sum, (1B) in effect states that given that M is the means to E, one has a reason (either not to intend E or to do M). If one does M one (either does not intend E or does M) and thereby one conforms to the reason stated in (1B) and to the ought-fact (Broome's term) stated in (1A). Hence doing M is an intelligible action. It is intelligible that people should take action to conform to reasons they have. That is indeed what it is for those reasons to be reasons. They make action taken to conform to them intelligible. So both undetachable conditional ought-propositions, and their corresponding propositions about conditional reasons, entail that one has a reason to take an action just because it is the means to one's end.

Broome is right to say that (1A) and (2A) do not justify the conclusion that one ought to do M. But the explanation of that has nothing to do with the relations between means and ends. Ought propositions state something like undefeated reasons. (1A) and (2A) do not yield the conclusion that one has an undefeated reason to do M. Even if M is the only means to E, doing M is not the only way of conforming to (1A) and to (2A). One could do so by not intending to do E. While (1A) and (2A) state (all the time assuming that M is the means to E) reasons of equal stringency for doing M and for not intending E, there may be other factors that discriminate between the two ways of conforming with (1A) and (2A), and that may show that even though one has a reason to M, one ought not to do so. The inference will be valid only if one adds to it a closure premise (e.g., that there is nothing else that bears on whether one ought to do M).

Hence the fact that (3A) does not follow from (1A) and (2A) does not support Broome's view (about the normative significance of having ends), whereas the fact that (3B) follows from both sets of premises shows his failure to achieve the goal explained by Wallace, that of explaining the normative constraints of instrumental rationality while denying that having an end is or provides a reason to pursue the means to its realization.¹⁸

As mentioned above, Broome points out that so-called theoretical reasoning is affected by an analogous problem: If P entails Q, it does not follow that if I believe that P I ought to, or have reason to, believe that Q.

¹⁸ In all these regards this case parallels the considerations that apply to facilitative reasons. I may have a duty to be in Oxford by noon, and buying a train ticket will facilitate getting there. But it does not follow from these facts alone that I have a duty to buy a train ticket, only that I have a reason to do so. The fact that there may be alternative plans facilitating fulfilment of my duty, that others may be preferable, etc., may explain this diminution in stringency. But my objection (in the text) to Broome's position does not rely on the facilitative principle. That deals with the relations between facilitative actions and the source action, whereas the objection to Broome concerns the relations between the source action and actions that constitute performing it (i.e. on the fact that making true one disjunct makes true the disjunction).

Among other reasons, this is to avoid the absurdity that since any proposition entails itself if I believe a proposition then I have reason to believe it (or ought to believe it). Here, too, we are looking for a way of explaining what is the fault with a person who believes the premises and the negation of the conclusion of a valid argument, without taking his belief in the premises as a reason to believe the conclusion.

Here, too, Broome's solution is to use undetachable conditional ought-propositions to dissolve the difficulty: One ought (if one believes the premises to believe the conclusion). Once again, however, it follows from this conditional that one has a reason to believe the conclusion, just as one has a reason to suspend belief in the premises. Either would constitute conformity with the conditional ought-fact. Again, until further features are drawn into the explanation, this is tantamount to taking the very fact of believing the premises as providing a reason to believe the conclusion, which is what we wanted to avoid.

Wallace adopts Broome's view on the matter:

I submit that we do well to interpret the instrumental principle along similar lines, as a constraint on combinations of attitudes that does not license detached normative judgements to the effect that we have reason to take the necessary means to our ends. Thus if you intend to do *x*, and believe that you can do *x* only if you do *y*, then the instrumental principle imposes a normative constraint on your attitudes. You can comply with this constraint either by giving up the intention to do *x*, or by forming the intention to do *y*. But it does not follow from the constraint, together with the fact that you intend to do *x* and believe that you can do *x* only if you do *y*, that you ought to intend to do *y*. (*ibid.*, 17)

Unfortunately, while (for the reasons explained above) it does not follow that you *ought* to intend to do *y*, Wallace does not explain what blocks the conclusion that you have *reason* to form the intention. Wallace himself notes that you can comply with the constraint by forming the intention. Other things being equal, that is enough to show that you have reason to form the intention. But that, as Wallace says, is unacceptable.

An alternative approach has been offered by Michael Bratman. He, too, argued that one's ends and intentions in themselves are not, and do not provide, reasons for action. He was therefore faced with the same problem, namely, how to understand and justify the view that, under some conditions, failure to follow one's ends is some sort of a fault, sometimes amounting to irrational conduct, or to having irrational attitudes. His reply is that the fault is in the process.

In assessing the rationality of an agent for some intention or intentional action, our concern is to determine the extent to which the agent has come up to relevant standards of rational agency. A failure on the agent's part to come up to such standards makes this agent guilty of a form of [criticisable] irrationality. In reaching such assessment our concern is with the actual processes that led to the intention and action, and with the underlying habits, dispositions and patterns of thinking

and reasoning [that] are manifested in those processes. Our concern is with the extent to which these processes – and the underlying habits, dispositions and patterns they manifest – come up to appropriate standards of rationality.¹⁹

The fallacy to avoid is the thought that irrationality in thought or action occurs only if one fails to conform to a reason (or a valid principle). It can consist in *faulty functioning*, that is in ways of thinking and of forming beliefs or intentions, and so on, which do not conform to standards of rationality.²⁰ People's ends may be relevant because they are the starting or initiating points of the processes whose rationality is in question. Failure to take advantage of opportunities to pursue one's ends can be *evidence* of irrationality in the agent, in the mental processes that failed at least on this occasion to conform to rational standards of deliberation, of belief or intention formation, of coherence of belief and intention or others.²¹

I will refer to such standards as standards of deliberative rationality. They include much more than standards governing the capacity to discern inferential relations, or their absence. They involve, for example, standards by which we judge the proper functioning of abilities to end deliberation when appropriate (and avoid the vices of dithering and indecision), abilities to stick with a conclusion (and avoid the vices of continually changing one's mind, feeling that the grass is always greener on the other side, etc.), as well as the ability to re-examine one's conclusions and intentions when appropriate (and avoid dogmatism, pig-headedness, etc.).

Much, though – as noted in the previous section – not all, of the discussion of instrumental reasons and of instrumental rationality aims to explain the very same phenomena that I identify here as a certain aspect of our rational functioning. I will here confine the term “instrumental irrationality” to the malfunctioning of our capacity to react properly to perceived reasons that manifests itself in failure to pursue available means to our ends. A person may be irrational in this way only on a single occasion, but, even then, his failure manifests a malfunctioning capacity, though one that may function properly most of the time.

Nothing has been said so far to explain why having ends affects the way we should think and form intentions, but, without such explanation, it is not clear how having ends is relevant to the well-functioning of our

¹⁹ *Ibid.*, 51. Other differences in their views notwithstanding, a broadly similar attitude on this aspect of the nature of rationality is taken, among others, by R. Nozick in *The Nature Of Rationality* (Princeton U.P. 1993) and by P. Railton in various papers, e.g. “How to engage reason: the problem of regress” in *Reason And Value* edited by R.J. Wallace, P. Pettit, S. Scheffler, M. Smith (Oxford U.P. 2004) p. 176.

²⁰ See my remarks on the subject in *Engaging Reason* (Oxford U.P. 1999) Ch. 4.

²¹ Not all such failures will be relevant to assessment of how rational people are. See a brief comment on the issue in Section 4 below.

deliberative processes. One explanation is offered by Wallace. He is one of those who trace the normative force of the instrumental principle to the incoherence of the beliefs that its violation involves. He argues that (1) if we have ends and do not intend to take what we believe to be the necessary means toward their realization, we have contradictory beliefs, and that (2) that is the rationale of the normative constraint to which we are subject once we have ends.

Wallace's defense of the first contention assumes that one can intend to perform an action only if one believes that its performance is possible. Hence, if you believe that doing Y is necessary to doing X, which you intend to do, and you believe that you will do Y only if you intend to do Y, and all the same you do not intend to do Y then

You will be left in effect with the following incoherent set of beliefs (assuming you are minimally self-aware): the belief that it is possible that you do x, the belief that it is possible that you do x only if you also intend to do y, and the belief that you do not intend to do y. The incoherence of these beliefs is a straightforward function of the logical relations among their contents, suggesting that the normative force of the instrumental principle can be traced to independent rational constraints on your beliefs – in particular, to constraints on certain combinations of beliefs. (p. 21)

He follows up with a complex defense and elaboration of this argument to meet various objections, and to apply the argument to a wider range of cases, but I will not consider those. I will assume that the above is broadly sound and that instrumental irrationality involves agents in some form of conflicting beliefs or beliefs and intentions that establishes or makes it likely that some of their beliefs are false or that their intentions are not well-supported.

The question remains whether the explanation of instrumental irrationality is where Wallace locates it, namely in the fact that instrumental irrationality leads to holding this incoherent set of beliefs. It is natural to expect that instrumental irrationality involves violation of some constraint and therefore leads us into trouble. But did Wallace identify the source of the trouble or one of its consequences, or symptoms? In the next section I will say something about the kind of defect contradictory beliefs involve. To anticipate: The fact that a person has contradictory beliefs shows that at least one of that person's beliefs is false. The information is often of little value, as often it does not help identify which belief is false. In the present case we know which belief is false: According to Wallace, the instrumentally irrational believe that an action that they intend, and which in the circumstances cannot occur, can occur. What is wrong with that? Well, it is the having of a false belief, and that is undesirable. Let us accept with Wallace that the instrumentally irrational are always guilty of that falsity, and assume that they are guilty of no other. The question is how undesirable is the having of *that* false belief? Is it undesirable in a way that explains why the so-called principle

of instrumental rationality is one of the standards that determine well-functioning deliberative processes?

The answer is not obvious. When people who are not instrumentally irrational intend to perform an action that in fact they cannot perform, they may be led to spend much effort trying to bring about what is impossible. But the instrumentally irrational has no intention to take steps to facilitate the impossible action. That is supposed to be his problem. Now it appears to be his saving grace. These reflections make me suspect that, whatever the merits of his argument, Wallace has not told the whole story – that there is something more to say about the kind of failure that makes this kind of malfunctioning a form of irrationality.²²

One part of the missing story seems fairly obvious: If you are prone to instrumental irrationality, you are less likely to achieve your ends, whatever they are. That is what is hinted at by calling instrumental rationality a form of skill (as Wallace, following Aristotle, does). The term is not altogether appropriate here, as skills are more specific. Abilities to see, speak or think are not normally thought of as skills. Our reliance on them is too pervasive, unlike our skills as drivers, carpenters, or computer programmers. Still, the point is well-taken: These capacities are valuable primarily, or at any rate in large part, for the uses they can be put to. They are valued for making us more effective as agents.

Every answer invites a further question. One that is particularly pressing here concerns the normativity of instrumental rationality. If there is no reason to pursue our ends as such, why is there a fault of any kind in failing to do so? The answer is to remind ourselves what this kind of irrationality consists in: It is not about the end result. A situation in which we do not pursue the means to our ends may be better than a situation in which we do. It may be better morally, and it may be better for our well-being or for our interests, or in other ways. That judgment depends on the content of the ends, and of whatever there is to be said for and against the means on grounds other than their relation to the ends. The fault is not in the end-result but in the process leading to it. It reveals an executive defect in the agent, a limitation on his ability to be an effective agent.

What is the value of being an effective agent? Necessarily anyone who intends to perform an action or tries to perform it intends to perform it successfully or tries to perform it successfully. Agents may, on occasion, hope that they will fail, but while they may merely pretend to intend or to try, when they actually do intend or try, they intend or try to succeed. Being an effective agent is not a good separate from that of being an agent. To be an

²² For detailed and incisive criticism of Wallace's position see J. Brunero's, 'Two Approaches to Instrumental Rationality and Belief Consistency' in *Journal of Ethics and Social Philosophy*, www.jesp.org (2005) Vol.1 no.1.

effective agent is to be successfully an agent. As to the value of rational agency, it is, as is the value of capacities generally, in the value of the option to use it, and in the value of its actual use. Its effective use does not guarantee that value, but is a condition for it, as the value of what we do accidentally cannot be to the credit of our rational agency. Hence if the power of rational agency is valuable, so is its effective exercise. Being a rational agent (that is an agent who can respond to reasons) is a fundamental aspect of being a person, and its value is part and parcel of the value of persons.

We see now why even if Wallace is right and, when one is instrumentally irrational one has a false belief – which is no doubt regrettable – that is not the explanation of instrumental irrationality. The source of our sense that something there is irrational is not that the irrational agent's beliefs or actions are undesirable. It is that he is not functioning properly. His irrationality is in his failure to conform to the standard of adequate agency. Possibly, in the circumstances of the case, the outcome is better than had he functioned well. But his failure of rationality is nevertheless a blemish. This is no consequentialist reasoning: It is not that it shows that he is likely to act badly some other time, though it may show that. Rather there is an ideal of rational agency, which that agent failed to reach on that occasion. If the results were beneficial, that was lucky. It was not due to him.²³

3. What price contradictions?

Rationality consists in part in proper functioning. People who fail to pursue the means to their ends display or manifest a form of malfunctioning criticisable as a form of irrationality. Does the well-functioning approach avoid what Broome sought unsuccessfully to avoid, namely the conclusion that having ends provides reasons for taking the means to their achievement? That is far from obvious. Let it be agreed that the primary locus of the sort

²³ Nomy Arpaly in her *Unprincipled Virtue* (Oxford U.P. 2004) points out that failure to pursue the means to one's end may be due to realization, below the level of consciousness, that one should not. This may indeed be so. Sometimes such realization may warrant saying that the person concerned has abandoned his end, even though he may not be aware of the fact, or at least not fully aware of it. If so, then no irrationality is involved. On other occasions the situation is different and the person is torn two ways. He is neither fully willing to abandon his end nor to pursue it, a conflict that may be psychologically understandable. In such cases, however sympathetic we are, the case is still one of irrationality manifested in this paralysis, or unresolved state. I will return to these observations below. Here I am merely pointing out that my discussion does not imply that one acts for a reason only if the action is preceded by or is caused by reasoning.

of rationality (or its failure) that the so-called instrumental principle is about is in the functioning of mental processes, initiated (or failing to be initiated) by the having of ends, and leading (or failing to lead) to consequential beliefs and intentions regarding facilitative actions. Is it not nevertheless the case that those processes have end products, and that those may be faulty if they result from a faulty process?

When people's mental processes fail to function properly with the result that they fail to take appropriate actions to facilitate realization of their ends, they are left with inconsistent – or at any rate incoherent – sets of beliefs, or beliefs and intentions. The simplest case is when people believe in the premises and in the negation of the conclusion of a valid deductive inference. They then hold inconsistent beliefs. If the inference is non-deductive, but valid, the beliefs may not be inconsistent but suffer from some form of incoherence.

The challenge facing anyone who believes that one's ends are not in themselves reasons, and do not provide reasons, is to explain how the well-functioning approach avoids the conclusion that when the process is faulty and its result is, therefore, faulty, one has reason to repair that result. Or in plain language: How does this approach avoid the conclusion that, if I do not have the intention to take an action necessary to achieve my end, then I have reason to abandon my end, and also a reason to form the missing intention (though once I do one of these things, the reason for the other lapses)?²⁴

The answer is that there is no reason to avoid a contradiction as such. In extreme and rare circumstances, people who profess simultaneous belief in propositions that are straightforwardly contradictory, or contradict each other, have no belief at all (if now I profess to believe that John is alive and that John is dead, then I have no belief on the matter). More familiar are the occasions when the fact that people profess belief in fairly obviously contradictory propositions indicates that they do not have a stable belief, that they vacillate one way and another, etc. Most occasions when we hold contradictory beliefs are not like that. We can have stable contradictory beliefs either without being aware of the fact, or while being aware of it.

²⁴ The problem is most clearly recognized by Scanlon. He too appears to hope to avoid the conclusion that agents' goals are reasons for them to take the means to their realization. But he thinks that (at least in certain circumstances) agents should take their goals as reasons. The escape appears to be that that does not mean that they really are reasons, only that the agents are irrational if they do not take them to be reasons. But he also thinks that they have a second order reason to take their goals as reasons in order to avoid irrationality. See his "Reasons: a puzzling duality," Wallace, Pettit, Scheffler and Smith (eds.) *Reason and Value* (Oxford U.P. 2004) 231, at 236-7. I will argue that agents' goals are neither reasons nor (with the exceptions explained in the next section) should they be taken as reasons. Agents do not have a reason so to take them.

When we learn that there is a contradiction among our beliefs we learn (1) that some of our beliefs are false, and (2) that we hold some beliefs that if used together as premises in an argument may lead us astray in a special way. Big deal! We hope that we all know that some of our beliefs are false anyway. And the risk that we will actually be led astray not by the logical implications of our false beliefs, but by their contradictory features, is, for all practical purposes, negligible.²⁵

If we know not only that some of our beliefs are contradictory, but also that a specific subset of all our beliefs contains contradictory beliefs, our knowledge may be somewhat more useful. The general knowledge that some of our beliefs are false is so useless because it does not help us locate them. Knowing that a relatively small set of our beliefs contains contradictory beliefs may be helpful in locating the false one. Even if we do not know which belief is false, circumstances may indicate that we should suspend belief in all the propositions in the contradiction-infected set. But nothing follows about what we ought to do or believe and when we should suspend belief from the mere knowledge that a set of beliefs contains a contradiction. In particular the mere fact that, say, three propositions form a contradictory set does not mean that we should suspend belief in all of them until we find the false one. The logical paradoxes, for example the sorites paradox, teach us at least that.

Hence while the irrationality of the ways in which we do or fail to reason from ends to means and to other facilitating actions may land us in a contradiction, it does not follow that we have any reason to do anything as a result. In particular, we have no reason to form intentions to follow the means to our ends, nor to believe the conclusions of valid inferences in whose premises we believe to avoid the relevant contradictions in these cases. Naturally, Wallace does not suggest which of the contradictory beliefs that he alleges that the instrumentally irrational is landed with should be abandoned. And, in fact, there is no general reason to abandon any of them, or any combination of them.²⁶

To conclude: There is nothing wrong with holding contradictory beliefs as such, and the fact that one does is no reason to change one's

²⁵ These remarks merely reiterate the importance of the difference between inferential relations among propositions or a person's beliefs and the dynamic process of belief adjustment and revision that is pointed out in G. Harman, *Change in View: Principles of Reasoning* (Cambridge, Massachusetts; M.I.T. Press/Bradford Books: 1986) and others.

²⁶ It may be tempting to think that as the intention to pursue the end is formed first, the intention to pursue the means, and the beliefs attending its formation, should be the ones to be adjusted, or abandoned. But Arpaly's argument shows this not to be so. In some cases the reluctance to intend the means is evidence that there is good reason to abandon the end

beliefs. At most we could say that we should abandon our false beliefs. But that is so not because of the contradiction. Knowing that a set of propositions is contradictory has epistemic relevance: It tells us that the contradictory set contains a falsehood. It may be part of a case for believing that one particular proposition is false. But it is no such case by itself. Without such a case we have no reason to abandon any of them. For all we know, we may then abandon a true belief and remain with false ones. Nor do we have reason to suspend belief in all the propositions in the contradictory set. The cost, epistemic and otherwise, of doing so may be too great.²⁷ That is why the logical paradoxes are rightly not generally taken as a reason to suspend our acceptance of the principles that generate them.

If principles of logic apply not only to propositions but also to intentions, then the preceding remarks apply to intentions too, as they do to other conflicts that cast doubt on the truth of the conflicting propositions, but fall short of contradiction. That is why the well-functioning approach does not yield the conclusion that our ends generate or provide reasons after all.

4. The difference ends make – another round

I have tried to do justice to the view that ends affect whether those who have them act rationally without taking them to provide reasons, and to do justice to the view that there are facilitative reasons without connecting them to ends, or to a specific form of rationality. Yet, I believe, there is another way in which ends affect our reasons, and thereby the rational constraints we are subject to.

Start with the obvious: We lead our life by considering our options and choosing. Often we have a number of options the case for any of which is neither better nor worse than that for some of the others. We choose in a way that manifests tastes and other dispositions that are not themselves reasons for one or the other of the options. Many, and typically the more important, of our choices are pervasive (affecting many aspects of our life) and long range (meant to be pursued over long periods of time). If we assume that having such ends makes no difference to the reasons we have, then even if we choose one and set out to pursue it (say, that we choose a hobby, such as stamp collecting, or playing all of Beethoven's piano sonatas)

²⁷ I realize that these remarks raise many questions that I cannot but avoid here: Why not say that knowing (through the discovery of a contradiction or otherwise) that at least one of our beliefs is false is a reason to suspend all of them, though not an adequate one. Do I assume that we can decide, at least in some circumstances, whether to suspend belief or not? These and others cannot be dealt with here. However, I believe that the substance of my remarks applies whatever position we take on these additional issues.

the reasons we have to engage in some activity that contributes to our pursuit (say adding some rare stamps to our collection) are, other things being equal, neither better nor worse than reasons we have to engage in a different activity that belongs with another pursuit, one of those we did not choose.²⁸

We may think that given our initial choice we will be disposed to engage in activities belonging with it, and will not frustrate its pursuit by choosing activities that do not, regardless of the fact that they are no less attractive. Moreover, sometimes the actions we took in the past incurred costs of various kinds that make deviation from our previously chosen path no longer rational. The alternatives have become less attractive. We may have actually spent money on equipment, say golf clubs, etc., whereas were we now to start on tennis we would need further expenses, a fact that makes it less rational to do so. Or we may just have delayed starting on the alternative activity, which means that we are less able to reach as high a level of accomplishment in it as in the option on which we did already make a start, and so on. These factors, and the fact that there is nothing inherently wrong with changing course and abandoning our initial choices, may be thought to show that the account so far given needs no supplementation.

I think, however, that we still miss something. One additional aspect of the situation is already covered by the earlier discussion of instrumental rationality. While there is nothing inherently wrong in changing course, a tendency to do so every other day shows indecisiveness, dissipation, lack of purpose or other mental dispositions that make one a less effective agent, and, since they relate to one's response to perceived reasons, reflect on one's rationality.²⁹ This, however, does not account for an important feature of our experience: We rely on the goals we have as contributing to the reasons to choose activities that belong with them, rather than those no less attractive activities that do not. How are we to understand this phenomenon? First, it does not involve changing our goals. I have in mind choices when this is not on the horizon: We feel no temptation or inclination to do so, and have no conclusive reason to do so. We choose among alternatives, and we notice that one of them belongs naturally with one of our goals (e.g., being an opportunity to improve, display or contemplate our stamp collection) whereas the others do not. Some actions that belong within a goal have no point (unless they happen to facilitate something worth doing) if one does not have it. Buying collectable stamps is probably an example. Other actions may have value in themselves, but their meaning to agents who have the goal is different than to those who do not.

²⁸ This applies to facilitative activities, but it is more helpful here to focus on activities that constitute pursuit of one or another possible ends.

²⁹ That is, very roughly, where Bratman leaves matters. I believe that in doing so he ignores the points made below.

Perhaps the best way to represent the way our goals affect our reasons is to take them to activate conditional reasons that we have anyway, though sometimes the condition affects not the very application of the reason but its strength or stringency. Actions whose point depends on (or is affected by) being taken as part of the pursuit of some goal are ones that not everyone has reason to take. There is reason to do so only for those who pursue the goal. If the goal affects only the significance of the action, then the strength or stringency of the reason for the action is conditional on pursuit of the goal. For someone intent on running a marathon every day during August, running a marathon today, the 20th of August, is crucial to the realization of his ambition. For me it is just an opportunity to know what running a marathon feels like – a matter of much less moment. On other occasions the option may be more valuable when it is a stand-alone option than when it is part of a goal, being a unique opportunity of some interest to those who do not share the goal, while rather a routine one for those who pursue it.

Crucially, the way goals acquire their normative relevance is by being conditions on the applicability or stringency of reasons. Therefore, they can have that effect only if the goals are worth pursuing in the first place. On this point my account is close to Korsgaard's.³⁰ It combines two elements: the normative impact of a potential goal depends (a) on its being worthwhile, that is, on there being reasons to adopt the goal, and (b) on the contingent fact that it is actually the agent's goal. If either is missing then the conditional reason does not apply.

These points do not affect the view of facilitative reasons and instrumental rationality proposed in this essay, but the rational impact that our goals have had to be mentioned to avoid the misleading impression that the previous discussion of ends would have left had it been taken to imply that they have no other rational relevance.

5. The Myth of Instrumental Rationality

I can no longer ignore the strangeness of the view I am defending. Professing to explain instrumental reason and instrumental rationality, I have drawn a distinction between facilitative reasons that are, I suggested, the reasons that people have in mind in those areas of thought which philosophers analyze as having to do with instrumental reasons, and the form of rationality which philosophers normally call instrumental rationality. Indeed, I drove a wedge between them. Instrumental rationality, it turns out, is not a matter of conforming to facilitative reasons. I followed Bratman in

³⁰ *Op. Cit.* It does not of course endorse her anti-realist way of thinking. Interestingly, however, much of what I suggest here can be readily adapted to non-realist frameworks.

claiming that instrumental rationality consists in the proper functioning of some of the mental processes leading to formation of beliefs and intentions. And I followed those, like Bratman, Broome and Wallace, who deny that ends or intentions as such constitute or provide reasons. At the same time, facilitative reasons, which explain our view that there are instrumental reasons, have nothing to do with the ends that people actually have. What explains this divide between facilitative reasons and instrumental rationality?

The answer is that there is no *distinctive* form of rationality or of normativity that merits the name instrumental rationality or normativity. In particular there is no specific form of rationality or of normativity that concerns the relations between means and ends. Philosophers fostered a myth of instrumental rationality, sometimes taking it to be the only, sometimes the simplest and clearest type of practical rationality or of normativity. In doing so they created a hybrid that lacks unity as much as it lacks distinctness. The discussion so far explained the lack of unity, the absence of any special relations between facilitative reasons and so-called instrumental rationality. It is time to turn to the question of the distinctness of each of these subtypes of instrumental normativity.

While facilitative reasons depend on source reasons, in that they are the source reasons plus the facilitative relationship (which is typically a non-normative relationship) between the facilitative action and the source action, reasoning about whether we should take a facilitative action is ordinary practical reasoning. It depends on what other reasons (facilitative or other) there are for that facilitative action, and what reasons (facilitative or other) there are against it, a reasoning that includes considering available alternative actions and the case for and against them. I do not think that there is any distinctive form of reasoning concerning facilitative reasons that can be called instrumental reasoning and underpin a form of rationality that can be called instrumental rationality. The facilitative character of some reasons is important since it may affect the case for the actions they are reasons for. Facilitative reasons obey the implications of the facilitative principle, which means that they have some properties that distinguish them from other reasons. But there are many other kinds of reasons, all with some distinctive features, as well as many common ones. Whenever reasons both (a) can bear on the case for and against one and the same action or attitude, and (b) can conflict, they can feature inseparably in the same arguments, and in the same deliberative processes, and therefore share the same form of practical rationality.

That form of rationality is the one that directs us in determining what we have most or undefeated reason to do, or which attitude is supported by the best or by undefeated reasons. Irrationality is manifested in failing to realize what those reasons are or to conform with them, under certain conditions (when the agent could have known what they are, etc.). There is

nothing that can be called instrumental rationality there, not even where facilitative reasons are involved.

Are things any different when we turn to the way that having ends affects our mental processes? The cases normally identified as manifesting instrumental irrationality are failures to conform to certain deliberative standards. This raises a major worry: Is there anything distinctive to instrumental rationality or irrationality, so understood? If it is just a matter of proper functioning of some of our executive abilities and skills, a matter of effective agency, is it a form of normativity failure that can be regarded as failure of rationality? Clearly, not every failure that makes us likely to fail to realize our ends is a failure of rationality. For example, some people tend to fumble, drop things, mix them up, mix up left and right, etc. None of these makes them instrumentally irrational, however much they affect their abilities to realize their ends. What distinguishes instrumental irrationality from those other failures?

Broadly speaking the answer is obvious: Instrumental irrationality is failure of rational capacities that has a specific effect on the ability to realise one's ends. But what are one's rational capacities? Again there is a ready broad answer: They are those capacities that are involved in discerning which features in the world merit a response, and how to respond to them, including both intellectual and motivational capacities (such as deliberative capacities, ability to come to a conclusion and to stick to it).³¹ The distance between such a broad and abstract answer and the detailed specification of the rational capacities is great, and that road cannot be followed here, and is not specifically relevant to instrumental capacities, if there are such. It affects our understanding of rational capacities in general.

Even if these observations are along the right lines in explaining why so called "instrumental rationality" relates to the degree to which our deliberative processes conform with normative standards, deviations from which manifest failures of rationality, they leave unresolved the worry about the distinctness of so-called instrumental rationality. Is it not the case that the rational standards that should govern our deliberative processes, which are or should be initiated by the adoption or the having of ends, are simply the same standards that should govern all our deliberative processes?

Take one example to illustrate the point: Sometimes the explanation for failure to take the means to our ends is weakness of the will. We know (or believe we do) what we ought to do, but are affected by various forms of self-indulgence that overcome our better judgment and we act against it.

³¹ This establishes that the position supported above conforms to the broad view of rationality as dependent on normativity, or, as I put it elsewhere, as dependent on our capacity to appreciate reasons and respond to them. See, *Engaging Reasons* (Oxford U.P. 1999), Ch. 3.

Now, weakness of the will is not confined to cases where we fail to take the means to our ends (say fail to engage in the vigorous exercises which are essential to our health, whose fate is dear to our heart). Weakness of the will can make us fail to adopt ends that we know (or believe we do) that we should adopt. If I believe that I ought to care about how well I teach my students, but I do not, I display weakness in adopting ends. If I do care about how well I teach my students, but I fail to give them detailed comments on their essays, in spite of my belief that commenting on their essays is essential to teaching them well, I again display weakness, which this time is a manifestation of instrumental irrationality. It seems plausible, however, that the standards by which I fail are the same in both cases.

Not all cases of so-called instrumental irrationality are cases of weakness of the will. They may involve other failures of rationality. To give but one other example, chronic dithering and indecision may make me fail to take the means to my ends. My dithering may make me miss the opportunity to realize my ends. But clearly chronic dithering and indecision may also make me fail to adopt ends that I am in a position to know that I should adopt. I cannot bring myself to conclude what is obvious: that this is an end to adopt. That, too, is a form of irrationality, this time about ends, not about the choice of means to existing ends. Again, the source of the failure is the same, regardless of whether it affects choice of ends or of means.³²

It appears that there is no such thing as instrumental irrationality. That is, there is no distinctive set of deliberative standards that are involved in getting us to reason correctly from ends we have to means, and that are different from those that are involved in reasoning about which ends to have. Of course, there is a difference between facilitative reasons and others. Facilitative reasons have a special kind of dependence on source reasons. But that is a difference in the content of our deliberation, not in the standards that should govern the deliberative processes. They are the same when we try to determine our will to adopt, or maintain, or abandon some ends, as when we try to determine what facilitative steps to take in pursuit of goods we take ourselves to be pursuing as our ends.

If so, why the widespread view that instrumental rationality is a distinctive and particularly unproblematic form of rationality? The answer is well known. Those who find it difficult to find a place for normativity (for values and reasons) in the world in which, they think, everything can be

³² The fact that failure to adopt the available means to one's ends does not result from failure to conform to distinctive standards; the fact that it is failure to conform to the general standards of deliberative rationality that merely manifests itself in failure to pursue means to one's ends, helps with the question with which section 3 ends, namely: What is so bad about instrumental irrationality?

explained by the physical sciences, are tempted to focus on so-called instrumental rationality for at least one of the following three reasons. Some think that normativity can and should be explained away in favor of accounts relying on motivations, subjective preferences, coherence relations among one's attitudes or the like.³³ Others, conceding the "objectivity" of normative relations and properties, think that instrumental rationality can be reconciled with their world view because it is all there is to practical normativity, and instrumental rationality itself can be reduced to some analytic truths about the relations between having means and having ends, perhaps with the additional admission that reasons for belief also are involved, they being less suspect than reasons for action. Finally, there is the common view that animals of some other species do have goals and direct their behaviour toward them, that is that they have the capacity for instrumental rationality, even though they do not adopt their ends for reasons, and they lack the capacity to deliberate about their ends. This seems to imply that instrumental rationality is a separable and more basic type of rationality.

Ends are, or can be, psychologically complex. At one extreme to have something as one's end is nothing but intending to do or achieve it, immediately or in the near future. But the notion is most at home with more remote and complex states, which consist in and manifest themselves in a range of attitudes, dispositions, beliefs and actions, over a period of time. This complexity allows for various forms of mismatch between the dispositions, intentions, attitudes and beliefs. Some mismatches will simply show that the person, or other animal, no longer has the end. Others will establish the existence of some deviations from the normal for those who have ends, which fall short of abandoning the ends. They can be attributed to stupidity, incompetence, or, sometimes, irrationality.

The fact that the pursuit of the means is part of what makes one have the ends misled some into thinking that instrumental irrationality is impossible. If one does not pursue what one believes to be the necessary and available means, then one does not have the end. With a simple end this may be close to the truth. If I do not reach for the glass on my desk, to sip from it, then can it be my end to drink its content now? However, even in such cases, it is not failure to pursue the means alone, but also the absence of feelings of conflict (I should really reach out for the glass. Why don't I do so?, etc.), guilt and remorse that establish that I do not have the end. Most ends are not that simple, and their attribution depends on various criteria. Conflict among these criteria does not always establish that the person does not have the end. Hence the possibility that some mismatches are due to failure of rationality.

³³ Among those who challenged such views is J. Hampton *The Authority of Reason*, (Cambridge U.P. 1998), Chs. 4 and 5.

Since ends range from simple to complex, it is not surprising that the criteria determining their existence, the criteria for the attribution of ends, vary between animal species. Since the mental abilities of some species do not enable them to engage in complex hypothetical thinking, or to have any grasp of distant future prospects, their members can only have relatively simple ends. With many species, reasoning about alternative ends is impossible. Hence they cannot adopt ends for reasons. But they may still have ends and direct their conduct and attitude in light of those ends, set themselves to achieve these ends, consider alternatives, try, fail and try again, and so on.

The existence of behaviour in pursuit of ends in creatures that cannot adopt ends for reasons encourages the thought that the same is true of humans. The refutation of that view is not a matter for this essay. What is relevant here is that some of those who accept that there are reasons for ends, and that humans are capable of adopting and changing ends for reasons, take the existence of ends and the pursuit of means by creatures who cannot reason about ends as evidence that there are radically two different forms of rationality involved. It is, however, a mistake to think that animals of species with more extensive mental capacities simply have what the others have and more. Rather, the existence of more extensive mental capacities transforms the functioning of many of the more basic ones. Even the most biological rooted needs, such as eating, defecating, having sex and finding shelter from extremes of temperature, are, with humans, subject to deliberation leading to their endorsement, rejection, channelling and control in a variety of ways. Hence with creatures capable of reasoning about ends, reasoning about means is not distinctive and special, but part and parcel of our general rational functioning. The thought that, with humans, instrumental rationality is a distinctive and more readily established form of rationality is a myth, which may well owe its hold on us to the three factors mentioned above.