



INSTRUMENTAL RATIONALITY: A REPRISE

BY JOSEPH RAZ

JOURNAL OF ETHICS & SOCIAL PHILOSOPHY

SYMPOSIUM I | DECEMBER 2005

URL: WWW.JESP.ORG

COPYRIGHT © JOSEPH RAZ 2005

Instrumental Rationality: A Reprise By Joseph Raz¹

THE OPPORTUNITY TO REPLY to four challenging comments enables me not only to remove some misunderstandings, but also to develop some ideas regarding which “The Myth of Instrumental Rationality” (“The Myth,” as I shall refer to it) was silent or even misleading. First let me set out the main theses of The Myth.

Assuming, as I do, that facts about the value of kinds of action constitute reasons for or against them, and that “rationality” or “irrationality” (or degrees of either) refer either to the capacity to appreciate and respond appropriately to reasons, or (in other contexts) to the way it is exercised, I argued that:²

1) There are instrumental or as, for reasons given, I call them “facilitative” reasons: 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 which facilitate it.³

2) The rationality or lack of it displayed in our reactions to our own ends – in particular adopting what we believe are means to their realization, or our failure to adopt such means – consists in the exercise or failure to exercise properly the same capacity, and in conforming with or violating the same principles of rationality, which we exercise or fail to exercise, conform to or fail to conform to in other contexts. The supposition that there is a special type of rationality, or of rational principles, to which “instrumental rationality” refers is the myth of the title of my article.

The two theses provide an alternative to a familiar account of these matters, which is rejected in the third thesis:

3) The fact that one has an end does not provide reasons for its realization, nor to take the means for its realization.

¹ I am grateful to Ulrike Heuer and Andrei Marmor for comments on an earlier draft.

² These formulations of my assumptions and of the theses are rough and ready. I have provided at least some of the required qualifications and clarifications in The Myth and others elsewhere. Here they are presented in abbreviated form to remind the reader of the outline of my position. I use “provide reasons” to cover both facts which constitute reasons and facts in virtue of which the facts which constitute reasons constitute reasons: for example, that you promised provides a reason to take the promise-keeping action, and that the action is a promise-keeping one is the reason for taking it).

³ Schroeder repeatedly but mistakenly implies that I claim that there are no instrumental reasons. Nor do I contend, as he claims, that they do not have “anything to do with ends.” My argument is that *adopting or having* a goal is not and does not provide a reason, but I try to explain in Section Four the ways in which having (worthwhile) ends can affect the reasons we have. Sobel misunderstands the facilitative principle. Having neglected to notice that it relates to derivative reasons that we have because we have other reasons to do something else, he takes it to be a principle of much grander ambition, like a claim that reasons derive from values.

I will start with a defense of the first thesis against a criticism made by Broome, which will lead me to reaffirm my critique of his account of instrumental rationality. Most of my response improves the explanation of the second thesis: Section Two identifies the phenomena I used it to explain. Section Three explains an important feature of the operation of our rational capacity and its role in accounts of normativity. Section Four reaffirms the generality of standards of rationality. Section Five, about contradictions, completes the defense of the second thesis, and begins the defense of the third. Certain concerns which sometimes lead to rejecting the third thesis are in fact consistent with it. The first two theses are the main⁴ part of my argument that that is so. I felt that that part of the argument for the third thesis is new and dwelt exclusively on it. As Sobel and Schroeder reject the third thesis, I outline, in the final section, the structure of an argument for it.

1. Defending the first thesis: Broome, Reasons and Requirements

Broome used to think that one ought (and therefore has reason) to intend whatever one believes is a necessary means to one's ends.⁵ In *The Myth* I argued that if that is so then, contrary to Broome's own view, one has reason to intend actions one believes to be the means to one's end. I wrote (referring to the reason not to fail to intend the believed means to one's end):

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 [the end] and fails to do M [the believed means]. 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).

If that argument is right, then if one has reason to bring about a disjunctive state of affairs (i.e., if one has reason to see to it that P or Q is the case) then one has reason to bring about either of the disjuncts. This does not

⁴ There are others, e.g., pointing out that since agents believe that their goals are worth pursuing, they believe, when they have a goal, that they have reason to take facilitative actions (the mistake, if one is made, is only in identifying what the reason is).

⁵ In his contribution to this symposium, he writes that he no longer endorses this view. I will not discuss his new view. He continues to use the term "requirement." Earlier he defined "a requirement" so that (where "X" is a variable for persons and "P" and "Q" for states of affairs) "P requires X to see to it that Q" means the same as "X ought to see to it that (Not-P or Q)." That was a stipulative definition, and none the worse for that. However, Broome has now abandoned that definition and is using "requirement" in its natural sense, to refer (roughly speaking) to that which is required by someone or by some sound or unsound principle or system. Requirements need not be normative, i.e. reason-providing. Broome may think of the requirements of rationality as constitutive of it, as contributing to a definition of rationality. The next section and the rest of this article show that when understood in this way his requirements of instrumental rationality are incorrect.

mean of course that if one ought either to do A or to do B, then one has reason to do A. Such a disjunction may be true because one has a reason specific to just one of the actions mentioned. By performing one of those actions one may be doing something one has no reason at all to do.⁶

I therefore concluded that even though Broome endorses something like my third thesis, his own explanation of instrumental rationality is inconsistent with it. The inference pattern I relied on in this argument is fundamental to practical reasoning: People have reason to do what will bring them into conformity with reasons which apply to them.⁷ I also relied on it in arguing for the first thesis, for the validity of the facilitative principle.⁸

Broome thinks that the argument is inadequate and the conclusion false. I pointed out⁹ that the facilitative principle, while depending on the validity of this form of inference, cannot be used to support it. Broome agrees. However, in the course of discussing the facilitative principle, he adduces an example which he takes to refute the form of inference I use against him:

Suppose you have an undefeated reason to avoid feeling hungry this afternoon. There are two alternative plans by which you can facilitate that result. You can eat the tasty and nutritious lunch that is already set on the table in front of you, or you can kill yourself. According to the facilitating principle read with "each one" for "any one," it would follow that you have reason to kill yourself. But plainly it does not follow. Suppose you have some reason to stay alive, perhaps a very slight one. Then you ought to stay alive. The fact that killing yourself would be a way to avoid feeling hungry does not count in the least degree against your reason to stay alive. You have no reason to kill yourself.

Broome takes it to be evident beyond a need for argument that you have no reason at all, however weak, to kill yourself in order not to feel hungry. But remembering Schroeder's point that "if ... I say ... that there is a reason for you to do A, the presumption that I have a relatively weighty reason in mind will be reinforced," it may be worth examining the case more carefully. Possibly there is such a reason which (because of Schroeder's presumption) would not, in the circumstances, justify *saying* that there is a reason.

I already mentioned the argument which, unless defeated, would show that there is such a reason, namely that doing so appears to be a way of accomplishing what we have reason to accomplish. Is there anything which defeats this conclusion? It is easier to state what does not. That you have a

⁶ Broome seems to confuse the two kinds of disjunction when he attributes to me the inference pattern: "one has reason (to X or Y); therefore one has reason to X."

⁷ I have discussed the considerations which underlie this inference pattern in "Numbers, With and Without Contractualism," *Ratio* 16 (2003) 3.

⁸ Broome, thinking that form of argument to be patently invalid, charitably interprets the facilitative principle some other way. He should not have been charitable. The principle is committed to the view from whose folly he wanted to save me. It happens to be true.

⁹ The Myth, p. 14, n. 18.

reason, of whatever strength, not to A (including that there is a better way of achieving what A would achieve) is no argument that you have no reason to A. This is so because, special cases apart, the existence of a reason for an action is logically independent of the existence of reasons against it. Is there anything to show that one has no reason at all to kill oneself, other than that (often) there are strong reasons against killing oneself?¹⁰ One such argument may be that killing oneself is not a way of not feeling hungry, a state which presupposes being alive. But that clearly is not why Broome is using the example, and it is not a consideration which will help with many other examples (how about killing my aunt to get the fare for a bus ride home). I therefore find Broome's example unconvincing, and he provides no other reason to doubt the validity of the form of inference which I used. Unless such reasons can be found, the facilitative principle is immune to this objection, while Broome's own account of instrumental rationality is inconsistent with his acceptance of the third thesis.

2. The second thesis: when are we "instrumentally irrational"?

Among other things, The Myth's second thesis explains phenomena commonly said to display instrumental irrationality (without assuming that having ends provides reasons for the means). I did not, however, take care to note that the phenomena I have in mind do not coincide with those sometimes said to be instrumentally irrational. I felt that whereas the first thesis means that I differ from some other writers regarding what instrumental reasons we have, the second thesis challenges mainly the explanation of phenomena which are generally agreed to be irrational. I underestimated the implication of the first thesis. Since the facilitative reasons we have are not those people often assume, the incidences of inadequate and irrational responses to them are also not those often assumed. Thus I misled Broome, who writes: "At a minimum, rationality requires us to intend whatever we believe is a necessary means to an end that we intend. ... A person is necessarily irrational if she does not intend whatever she believes is a necessary means to an end that she intends. Raz agrees." Well, I do not quite agree.¹¹

First, and this point is made clear in The Myth, I do not agree with Broome if the implication is – as I believe Broome, along with Schroeder and Wallace, take it to be – that there is a specific principle of rationality applying

¹⁰ Is the example as rhetorically convincing when there is only 'a very slight reason' to stay alive? Suppose my only reason to stay alive is that it will enable me to water my beautiful flower (which pleases passers-by) this evening. It cannot survive more than two more days, but will die tomorrow if I do not water it tonight. Is it still clear that I have no reason at all to kill myself in order to avoid feeling hungry if I have no other reason not to live?

¹¹ Schroeder was also misled. I do not think that I said or am committed to holding, as Schroeder alleges, that "on Raz's account, cases in which an agent has an end and does not take the necessary means are all cases of improper cognitive function."

to this type of case, if it implies that there is something like “the instrumental principle” of which Wallace and Broome offer divergent analyses. The second thesis denies precisely that. It denies that there are principles of rationality specific to instrumental intentions, as I will call intentions to pursue the means, or more broadly to take facilitative actions, necessary or otherwise, to one’s ends. It claims that when failure to form or to have instrumental intentions is irrational, that is because it violates general principles of rationality, principles which are not about forming or having instrumental intentions.

This point raises the question whether every time one has an end and does not have a relevant instrumental intention one is irrational, having violated some principle of rationality. Here too I disagree with Broome, as well as with Wallace. In *The Myth* I did not raise this question, and misled the reader into thinking that I take the answer to be affirmative. But that would be a mistake. At a minimum one must allow for a time gap between adopting a goal and forming an intention to pursue any particular action or plan to bring about or to facilitate its realization. This is particularly obvious when there are various plans by which it may be realized, and one has to choose among them. In such circumstances some delay would be reasonable. I decided two weeks ago to visit a friend in Covent Garden tonight. I could go there by underground, by bus, by taxi or on foot, yet I have no instrumental intention, no intention which way to get there. The visit is still some two hours away, and as whichever way I go I do not need more than an hour, I need not think about the way to get there yet. This is not a trivial or minor point. We all have ends without any intentions regarding any means for their realization, and the time gap need not be brief. I may intend to buy a new house in five years time, have no relevant instrumental intentions, and yet be perfectly rational. I may intend to retire to the seaside in 20 years’ time, and again have no instrumental intentions, and be none the worse for it. Moreover, I may abandon such a goal after 10 years, never having had any instrumental intentions and having been innocent of any irrationality as a result.

It could be that all the feasible plans for some goal have a common step, which is therefore a necessary step for the realization of the goal. One may form an intention to take that necessary step, before any decision about which plan to pursue. But on the other hand, one may not form any such intention, and in circumstances in which it is not irrational not to have decided on the plan, it is not irrational not to have an intention to take that necessary step either. There is no reason to adopt that part of a plan before adopting the plan, even if it is known or believed that eventually one will be irrational if one fails to adopt it. Similarly, while one can, and philosophers do, regard the disjunction of the alternative plans as itself a plan (which, moreover, is or is believed to be necessary for the realization of one’s goal), there is no reason to adopt that plan before adopting one of the more specific plans which would facilitate one’s goal. Hence, in the circumstances illustrated in the examples, failure to intend the disjunctive plan, or some less

specific variant of it (such as the intention to do something about one's goal), is not irrational.

A delay in forming instrumental intentions may be reasonable for other reasons as well. Imagine: One of my goals is to give up smoking. My various attempts have so far failed. Once I decided to smoke one less cigarette every day. Having failed to do so, I adopted a different plan: first to limit the occasions I smoke (after meals, etc.), then gradually, week by week, reduce the number of permitted occasions. I failed in that too. I have not abandoned my goal, but I am yet to adopt a new plan for achieving it. At the moment I have no relevant instrumental intention. I think that I know what plan to adopt. I believe that I should simply stop smoking altogether, and that I should do so immediately upon adopting this plan. I should have a glass of water whenever the desire to smoke swells inside me. But I am still smoking. I have not yet formed the intention to follow that plan. I am still nerving myself up to do so. In such cases, and no doubt others too, I have a goal and do not intend any of the means to it, yet I am not – not yet – irrational.

This example shows that even if I believe that a particular plan will not only facilitate the goal, but also is the only way to achieve it, even if I believe that it is the necessary and sufficient means to the end, I may still not be irrational if I have not formed the intention to pursue it. My delay in forming an intention to try what I believe to be the only plan to achieve my goal may not be irrational. "Building up" one's resolve before taking the plunge can be useful, or even essential.

It may be thought that in examples like these, while no facilitative plan is yet adopted, the agent is not irrational, because he has an intention when and how to adopt a facilitative plan. Alternatively, it may be thought that in cases where there are alternative possible plans, the agent is not irrational because he intends to pursue the disjunctive plan — that is, he intends to pursue one or the other of the sufficient plans. Sometimes this is indeed the case. I may intend to adopt the plan once I feel strong enough to do so, or I may intend to wait till the weekend and decide whether to adopt it then. I may intend to decide which way to get to my friend an hour before our meeting, and so on. But, on the other hand, I may not. The situations I described are ones in which even if it may be rational to have instrumental intentions, it is not yet irrational not to have them. Rational agents form instrumental intentions sometime during such intervals. They do not need a prior intention in order to do so.

All the sketched examples illustrate situations in which it is not irrational to lack instrumental intentions. In cases like these it may be both rational to form such intentions, and rational not to do so. However, this need not always be so. It may actually be irrational to have instrumental intentions to pursue any particular means to one's ends (even when it is not irrational to have the ends). Obviously, not all instrumental intentions are rational. One can be irrational not only through failing to have instrumental intentions, but also by having such intentions that one is in position to know one should not

have. The kinds of examples I gave make it plausible to think that in some circumstances it would be premature to form any instrumental intention regarding one's (rational) ends. And if so, there will no doubt be circumstances in which it will be irrational for a person not to know that that is the case, and therefore also irrational to form instrumental intentions of any kind regarding one or another of his rational ends.

It would be good to subsume the preceding remarks, supported by examples, within a more abstract explanatory framework. I am able to offer only a partial analysis. I am relying on two general propositions. First, it is possible to believe that an action is best, or that one has conclusive reason to perform it without intending to perform it. Second, while we sometimes have intentions to form intentions they are not necessary to be able to form intentions (on pain of regress). We form intentions in response to reasons which, as we see it, make them intelligible or necessary. This is true not only of intentions to pursue ends, but also of intentions to pursue means. The thought that adopting an end involves forming an instrumental intention, if only an intention to take some means towards its realization, is prompted by belief that one can only account for the way having an intention sets one towards action by assuming the adoption of another – instrumental - intention. In fact it can also be accounted for by the fact that a rational being will form instrumental intentions when, as he sees things, it is appropriate or necessary.

Next, consider a possible objection to the facilitative principle. That principle indicates that we have many more instrumental reasons than is sometimes assumed. A mistaken chain of reasoning can lead one to think that, if we have all those facilitative reasons, then (assuming that a reason for an action is a reason to intend that action) we also have reasons to intend to take those facilitative actions. From this it is tempting to conclude that, when we have reasonable beliefs about worthwhile ends and about facilitative steps toward their realization, we are irrational if we fail to have such intentions. If this chain of reasoning were sound, it would refute the facilitative principle, for it would entail that we are irrational on many occasions when we are not.

There is, however, much wrong in this chain of reasoning. Relevant to our purpose is the misperception of the relations between reasons, including reasons to have intentions, and motivations, including intentions. I have discussed elsewhere¹², and cannot revisit here, a variety of ways in which too close a relationship between reasons, or belief in their existence, and motivation is often postulated. An example will have to do. Regarding each of the following three ends I have, and know I have, an undefeated reason to pursue it: have a holiday next spring on a Greek island; have a holiday next spring in the American Southwest; spend the time at home catching up on some writing. I should not realize more than one of these ends. At the mo-

¹² *Engaging Reason* (Oxford: OUP 2000), ch. 5.

ment I have no intention regarding any of them. I assume that it will be agreed that the circumstances may be such that I am not irrational or at fault in any other way as a result. Nor do I have instrumental intentions regarding any of these undefeated ends. Note that by the facilitative principle it may be rational to form an instrumental intention before forming an intention regarding the goal that it will facilitate: It may be rational to intend to take an option on a cottage in the Greek island in order to keep that option cheaper than it would otherwise be, even though I do not yet intend to have that holiday. But, equally, it is often rational (i.e. not irrational) not to form any instrumental intentions even while one has a reason to have such intentions, though possibly a time may come when such failure would be irrational.

In general, just as having (a reasonably believed or known) undefeated reason to have an end does not make failure to intend it irrational, so having (a reasonably believed or known) undefeated reason to have an instrumental intention does not make failure to have it irrational. To a degree, the point can be accommodated by distinguishing between reasons to have an intention and reasons to form it at a particular time (or when certain conditions obtain). Failure to incorporate such a distinction in their accounts of instrumental reasons explains some of the divergence between the incidences of irrationality assumed by my second thesis and those assumed by the accounts offered by Broome, Wallace and others.

However, I do not think that that distinction can completely explain the divergence between the phenomena identified as irrational by our respective accounts. It is unlikely that for all circumstances and regarding each option available in them which one has reason to take, there are conditions which determine a point (or a very short interval of time) before which one has conclusive reason not to intend it, and after which one has conclusive reason to intend it, let alone that such conditions can be known so that failure to conform to these reasons renders one irrational. Where there are no such conditions, there is a significant period during which one has a reason to take the action, and yet failure to form an intention to take it is not irrational. This will be true also where the reasons are due to the facilitative principle.

One issue I leave unexplored is the relationship between a reason to perform an act and a reason to intend to perform it: if there is now a reason for me to retire to the seaside in 20 years time, does it follow that there is now a reason for me to intend to retire, or do I have reason to have such an intention in, say, 19 years from now, or is it rather that I will then have a reason to form such an intention? The matter depends partly on the way reasons relate to time, and partly on the boundary between what reasons we have and how we function when we function rationally.

Dealing with analogous forms of reasoning¹³, some philosophers spot a mistake in not restricting rational pressure to the taking of means which are necessary for the end. When aware of two (and only two) appropriate ways of realizing whatever we believe that we have reason to realize, we are not irrational for failing to intend one of them. We are irrational only if we intend neither, for, as we believe, taking one or the other is necessary for the realization of the end. If that were so, it would mark another flaw in the chain of reasoning I have criticized above. However, the facilitative principle is about steps or plans which realize or facilitate realization of ends. Roughly speaking, it is about means “sufficient” to ends. This is another point of disagreement with all four contributors to the symposium. Schroeder calls the following “Schema Detach”:

“For all agents x and act-types m , if x has end E and m is a necessary means for x to accomplish E , then there is a reason for x to do m .”

But far from the question of its validity being “central to [my] argument,” as he says, I did not consider the schema at all. That was not because, as he implies, I “take it as obvious . . . and, as a constraint on an adequate account of instrumental reason,” that the schema is invalid. Rather, it was because the fact that some means are necessary to realize an end is irrelevant to the issues I discussed. True, *if having a goal is a reason to perform actions which constitute it or will lead to its realization or make it more likely*, then the *necessity* of an action for the realization of a realizable goal can affect the *stringency* of those reasons. The standing of means which are necessary to a goal one has is, however, irrelevant if having a goal is no reason to take means which facilitate its realization. In other words, the argument I advanced was against the thesis that having a goal is a reason for what we may crudely and inaccurately describe as means sufficient for its realization (or ones which increase its probability). The fact that some such means are also necessary for its realization becomes relevant only if the arguments I advanced in the article fail.¹⁴

3. The second thesis: rational functioning

There are two aspects to the second thesis. First, that one central sense of rationality has it as a capacity, and a related sense takes it to be the successful exercise of that capacity, that is its proper functioning, which cannot be accounted for entirely as a matter of conforming to reasons. Irrationality need not be acting against reason. It can be just failing to function properly (as determined by standards of rationality). Second, that there are no standards of rationality specific to dealing with believed means to goals one has. They govern such contexts, but that is simply because they govern all con-

¹³ Namely reasoning about instrumental reasons, but without assuming the facilitative principle.

¹⁴ See the argument regarding necessary means earlier in this section.

texts. Instrumental irrationality is simply irrationality manifesting itself in contexts of dealing with means to ends.

The first aspect explains how people who fail to intend the means to their ends can sometimes be irrational, even though having ends provides no reasons, and even when their ends lack any merit (and therefore the facilitative principle does not apply to them). Some of the contributors rejected my second thesis without addressing this aspect of it. Sobel appreciates the first aspect, but he distorts its meaning by thinking that it is a claim about “two sources of normativity.” I cannot say that my comments below will show that normativity has no sources, but I hope that they will show that the distinction I have in mind has nothing to do with different sources of normativity.

The general idea is that to act rationally one needs (a) to have reasons (which one can become aware of) and (b) to respond to them appropriately. These are two different categories: normative features of the world and the exercise of a capacity to respond to them. They cannot be reduced to one category, being radically different. Even though the capacity for rationality is norm-governed in being subject to standards of correct and incorrect exercise of the capacity, these norms are not like norms or principles for the guidance of conduct (principles like keeping promises).

So how is our rational capacity to be understood? I will try to clarify matters by considering Schroeder’s second objection to my claim that the value of a properly functioning rational capacity does not provide a reason for taking the means to our ends. The question is whether, where failure to take the means to an end one has is irrational, there is a reason to take the means to one’s end (as well as a reason to abandon the end) in order to avoid the irrationality. If so, then standards of rational functioning are simply reason-stating principles, like any other normative principle.

Sometimes people who fail even to try to take the means to their ends display or manifest a form of malfunctioning criticizable as a form of irrationality. Broadly speaking, that much is common ground. My argument was that what makes such failures irrational is not that the agent’s irrational beliefs or actions are undesirable, nor that he has any reason to avoid them, or their combination. It is that he is not functioning properly. There is, I said not very helpfully, an ideal of rational agency, which that agent failed to reach on that occasion [18]. It would have been better had I said instead that he failed by a standard of rational functioning.

Schroeder quite sensibly wonders (I express his objection in my own terms) how I can maintain both that the capacity for rationality is a constitutive part of our personhood, and therefore that it is valuable, and yet deny that we have a reason to avoid its malfunctioning on each occasion. Put in a rough and ready way, Schroeder’s argument is:

- (1) Having rational capacities is valuable; therefore
- (2) being irrational is, as such, bad (on each occasion); therefore
- (3) we have a reason to avoid being irrational (on each occasion); therefore

(4) (by the facilitative principle) we have reason to take the means to that end.
Each one of these transitions is unwarranted.

The value of a capacity does not entail that every case of its proper, or successful, use has value. This is very plain in the case of those capacities we exercise at will. It may be valuable to be able to swim or to play chess, or to drive, but we do not have reason to swim, drive or play chess whenever we can, nor whenever we could do so well – and not every time we do, not even every time we do so well, is our action of value at all. Sometimes there is nothing to be said for the use of the capacity, and sometimes it is better to do it badly, not only that on balance it is better to do it badly, but that there is no reason at all to do it well, and some reason to do it badly. Assuming that this would be agreed, I will not stop for examples. The explanation is that the value of the capacity is partly in the freedom it gives us, the choice whether to use it and when, and partly in the fact that sometimes its use is valuable. The two are interconnected, and the freedom would not be valuable if its use could not be valuable. But it does not follow that every use is valuable, or that we always have reason to use the capacity.

Rationality is not a capacity we use at will. Therefore a freedom to use it or not cannot be where its value resides. But the same structure applies to other capacities we use automatically, e.g. our perceptual capacities. So long as I am conscious I hear sounds, if, that is, I have the ability to hear, if I am not deaf. I can manipulate myself by blocking my ears, or trying to divert my attention, but these are very different from the decision to do something one has the capacity to do at will, like raise one's arm, or turn one's head. Hence, the value of perceptual capacities does not relate to freedom, but to their epistemic, aesthetic, sexual and social advantages. Still, while their value depends on the fact that some occasions when the capacity is used are valuable, it does not follow that all such occasions are, not even to the slightest degree.

Rationality is like perceptual capacities in being engaged willy nilly, in not being a matter of choice (though being subject to manipulation via the use of alcohol, etc.), but it differs from them in reaching deeper into who we are. As much as loss of sight, hearing or tactile sensations would affect me, it does not touch me in the way that complete loss of rationality, as in advanced dementia, does. Such loss means that I am no longer a person. The value of rationality (in the sense of rational capacities) is therefore different again. It lies in the value of personhood, of which it is a constitutive part. But yet again, while its value depends on its exercise being sometimes valuable, it does not imply at all that all its instances have value, not even to the slightest degree. Hence if, for example, one has murder as one's goal, there is no value at all in being rational about the pursuit of the means to that goal, and no

reason to take them, not even to avoid the irrationality of not taking them (while not abandoning the goal).¹⁵

The transition to the third proposition above is undermined by the fact that the use of our rational capacities does not require a decision or intention. Regarding such capacities, the question whether there is a reason (or motivation) to use them does not arise. We may have reasons to listen, but we cannot have reasons to hear. The case of our rational capacities is essentially similar, only more complex, and requires more distinctions which cannot be discussed here.

Examining the fault in the transition to the fourth proposition helps in bringing out the first aspect of the second thesis, namely that while rationality is a capacity to appreciate and respond to reasons, its exercise is not to be understood exclusively in terms of following or of failing to follow reasons. Rather, the exercise of rationality is a process, a functioning, which goes well or badly. We judge it by its success or failure to conform to standards which govern it. How does this affect the issue at hand? Schroeder's challenge depends on agents having reason to take the means to their ends, the reason being that that is one way of avoiding being irrational, or, if you like, one way of being rational. Suppose John intends (it is one of his ends) to get a bottle of milk (because his baby needs feeding). John knows that he will get one if and only if he goes to the shop. He has a reason to go to the shop, but it would be funny to think that he has a reason to go to the shop as a way of not being irrational on this occasion. Clearly the only reason he should recognize is that it is a way of getting the milk, which he has reason to do. It is true that, in the circumstance, if he does not, he will be irrational, and this is a case in which it is bad to be irrational. The point is that the rationality is the process which leads you to the reasons, and to the correct reaction to reasons; it is not itself a reason, it is not what the process is about.

Two final points: First, it is evident why I think that Broome is asking the wrong question when he asks whether there are reasons to conform to what he calls the requirements of rationality. He thinks of principles of rationality as reasons, or reason providing in the way one has reason to respect people. His question misunderstands what rationality and its standards are: that they are exercises of a capacity to respond to reasons rather than themselves reasons or reason providers. And it is a capacity whose exercise calls for no reasons, but is automatic as it were. Second, while rationality is exercised automatically, one may have reasons to remain rational or to end one's life or one's life as a rational being, and one may have reasons to manipulate one's exercise of rationality on occasion, or in general, just as one may have reason to improve one's hearing or reduce it, generally or on occasion.

¹⁵ We may still admire the skill of the murderer. But that is a point about a different dimension.

4. The second thesis: the generality of standards of rationality

So much for the first aspect of the second thesis. The second aspect denies that there are standards of rationality specific to instrumental reasoning. Broome and Wallace just ignore my argument to that effect. I have reinforced it here by the observations of Section Two, which show that both Wallace's instrumental principle and Broome's requirements concerning instrumental rationality are false. Given that it is not the case that it is irrational not to intend the believed means to one's ends, the considerations which determine when it is rational or irrational to intend those means or to refrain from doing so cannot derive simply from the fact that these are believed means to ends one has. This reinforces my argument in *The Myth* that they are considerations deriving from general principles of rationality.

Sobel thinks that this second aspect, that instrumental rationality is not a distinct type of rationality, rests on the view "that there is no general kind of normative mistake that we make in all and only cases in which we fail to take the acknowledged best means to our adopted ends." But that is not so. The claim is that there are no standards of rational functioning which are special to those contexts. I illustrate the case by examples of ways in which people are irrational when they do have ends, but irrationally fail to pursue them, and when they irrationally fail to adopt ends. As I read him, Sobel agrees with me that the very same standards of rationality govern all contexts of practical thought and conduct. His comment, though I am not sure that I entirely understand it, is that when people irrationally fail to adopt ends that they should adopt, they fail to follow the "recommendations" they make to themselves, and that is a failure to pursue the means to their own ends. I only know of recommending as an interpersonal activity, and as a metaphor it seems inappropriate to the example, which is one of dithering. One can recommend only if one knows one's mind, whereas one dithers only if one does not.

Sobel continues to cite my example: "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." [26] And observes "Raz should be curious if failing to adopt ends one knows or believes (Raz is incautious between these two formulations) one ought to adopt itself is an example of failing to take the means to one's ends." But there is no reason to think that chronic dithering and indecision occur only regarding the means to pursue one's intentions. The only reason which I can think of for accepting what Sobel says is that since (a) all irrationality is about the pursuit of means, and (b) this is irrational, therefore there must be an intention such that this irrationality is about pursuing the means to its realization. I suppose that this is not quite what Sobel thinks. Rather he thinks that

"a different, less dramatic, conclusion one might have reached is that instrumental rationality covers cases not only in which the relations between

ends and means are in play, but also cases in which we fail or succeed in heeding our recommendations about what ends we ought to adopt. Call such a view 'broad' instrumental rationality."

But this is not a different conclusion. This is just playing with words. The word game is to reason that since the standards of practical rationality are the same in all the domains of our practical life, why not call all of them standards of instrumental rationality. Sobel may protest that I ignore the role of "recommendation" which he takes to be analogous to intention and thus to legitimate his notion of broad instrumental rationality. But taking this thought seriously requires ignoring (a) that no recommendation takes place, that this is just a metaphor, and (b) that there is no case for thinking that ignoring recommendations is irrational and (c) that his use of the metaphor shows that Sobel does not appreciate the range of the psychological phenomena I discussed.¹⁶

5. The significance of contradictions

Schroeder rejects my third conclusion. He contends that having ends provides instrumental reasons and that my arguments to the contrary fail. Indeed my view of the nature of rationality itself shows that having ends constitutes or provides reasons for the means. His own case that ends provide reasons for means does not depend on failure to pursue believed means landing one with contradictory beliefs. But what I say about contradictions, he thinks, is incorrect and shows that I am committed to accepting that having ends is a reason for taking the means.

It is helpful to understand why I discuss contradictions: In *The Myth* I argued, roughly speaking, that even if failing to try to take facilitative actions one believes to be available involves having contradictory beliefs, that fact does not account for what has gone wrong, for what sometimes renders such failure irrational.¹⁷ I was non-committal regarding whether one has contradictory beliefs in such situations. Therefore¹⁸, just in case it turns out that one does, I sought to forestall an argument to the effect that as one has reason either to suspend belief in or to reject at least one proposition in any contradictory set, it follows (by the facilitative principle) that, if one has a goal, one has reason to pursue the means to it (as one way of avoiding the contradiction, the other being to abandon the goal). I argued that for all intents and purposes what is undesirable about contradictory beliefs is nothing more than that at least one of them is false. And while we may have a reason to avoid false beliefs because they are false, we do not have a reason to suspend

¹⁶ It is no more appropriate for weakness of the will, my other example, than for dithering.

¹⁷ This was part of my argument against Wallace's explanation of instrumental rationality.

¹⁸ And not because "He [namely I] assumes that if there is any ... reason [for thinking that having goals is a reason for the means], it must derive... from a reason not to have contradictory beliefs," as Schroeder writes.

belief in each of a set of contradictory beliefs because they belong to a contradictory set.

Schroeder finds this inadequate. He thinks that “it’s not at all obvious that believing something false is the only thing wrong with having contradictory beliefs.” I agree. That is why I spent some time to point out what reactions to the realization that one holds contradictory beliefs would be rational, inviting the reader to acknowledge that they depend on whether we can pinpoint the location of the contradiction sufficiently to justify suspending true beliefs in order to avoid a false one. That would tend to show that what we rightly care about is the avoidance of false beliefs rather than contradictions in themselves, or any other of their features. The contradictions are, I claimed, simply reasons for believing that we have some false beliefs. No doubt I should have better explained the point.

Unlike Schroeder, Sobel makes a concrete suggestion about what is wrong with having contradictory beliefs.¹⁹ He insists that “one could clearly hold that the problem with a manifest contradiction in one’s attitudes ... is not the undesirable upshot of such things but a more intrinsic assessment that to do these things is to be improperly functioning.” I do not know what doings he refers to. The Myth discusses contradictions, not manifest ones. A standard of rationality which requires, whenever the question of the case for any belief modification is raised, a complete examination of all one’s beliefs and their consequences for possible contradiction with the modified belief, is unreasonably demanding (requiring life to come to a complete standstill before any rational belief change). Besides, and this applies to manifest contradictions as well, The Myth gives reasons to think that sometimes avoiding a contradiction would be irrational, reasons not directly challenged by Sobel, but which undermine the claim that holding contradictory beliefs, even manifest ones, is a mark of a malfunctioning.

Wallace does not dispute my general contentions about the significance of contradictions. He rightly points out that in cases of failing to take steps one believes will facilitate realization of one’s end, one can avoid the contradiction (assuming there is one):

In the particular case at issue there is a special feature of at least one of these problematic beliefs that simplifies the process of revision in response to acknowledged local inconsistency. [One of the beliefs in the inconsistent] set is a belief about my own intentions; the truth to which it is answerable is a matter that is directly subject to my volitional control.

¹⁹ Perhaps I should mention here that The Myth dismisses as practically irrelevant the fear that contradictions may lead to accepting unsustainable inferences. Systems of paraconsistent logic enable one formally to control and eliminate that risk. Neither Sobel nor Schroeder disputes the point.

That is true. Contrary to what Wallace seems to imply, this does not mean that restoring consistency is easier²⁰, only that it can be achieved not only (as in most cases) by changing some of one's beliefs, but also by changing one's intentions, thus making some of one's beliefs true. But my criticism was not that it was difficult to restore consistency. It was that the existence of a contradiction is not a reason to do so. Wallace says nothing to reinforce his original suggestion that we have such a reason. He agrees with me that contradictions may be an epistemic reason, a reason to re-examine one's beliefs. But in the cases under consideration, no such examination is needed. If Wallace is right to think that there is a contradiction, then we know why. The question is whether *this* shows that there is anything wrong. My arguments seem to have convinced Wallace that this is not always the case, that sometimes avoiding contradictions is irrational. He says nothing to suggest either (a) that whenever we do not intend the believed means to our end we are irrational (i.e. that our functioning was defective), or (b) that there is a reason to modify either intentions or beliefs to avoid the contradiction. Section Two showed that neither is the case: it may be rational not to have instrumental intentions, and that need not be a reason to abandon one's end. Even where there is reason either to abandon the goal or to adopt an appropriate instrumental intention, that reason cannot be to avoid the contradiction, for – as I argued and Wallace does not dispute – contradictions do not provide reasons. Wallace mistakes my argument when he suggests that where the contradiction is local, my argument does not apply. It does. My conclusion was that the only reason present is to avoid a false belief. When people fail to take means which they believe to be necessary to their end, their false belief is known (at least as alleged by Wallace, but regardless of whether Wallace is right that they have contradictory beliefs). It is their belief that even as things stand, achieving their ends is possible.²¹

Wallace persists with his account for he believes that it explains the normative pressure we are subject to, and that we feel, to adopt instrumental intentions. This seems, however, not to support his account at all. The facilitative principle explains why we have that pressure: we believe that our ends are worth pursuing, and therefore, by that principle, that we should facilitate their realization. Wallace seems to doubt the adequacy of such an account on the ground that even akratic agents are susceptible to normative pressure to pursue the means to their ends. Akratic agents, however, act for a reason –

²⁰ Both beliefs and intentions are formed in response to reasons.

²¹ One can pursue the argument against Wallace's position beyond the point at which I left it in *The Myth*. I argued above that it is implausible to think that one has reason to form an intention in order to make one of one's beliefs true. Beyond that, it is arguably false that if one intends an action, one believes that it is possible rather than believes that it is possible if one performs appropriate facilitative actions. That belief is not false, however irrational one is in not intending the means.

the akratic guzzler of chocolate eats chocolate because he enjoys its taste, or whatever. In as much as he believes that he has a reason for the end, he also believes that he has a reason for the means which facilitate it. Of course, as he himself sees matters, he has an adequate reason neither for having the end nor for pursuing the means, but at some point rationalizing explanation cannot explain irrational conduct.

6. The third thesis: why having goals does not provide reasons

Schroeder thinks that there are powerful arguments against my third thesis, arguments which establish that having ends provides reasons to take the means. One argument by which he lays much store is that “[Schema] Detach postulates a clear asymmetry between taking the means to your ends, and giving the end up. And importantly, there should be some such asymmetry,” by which he presumably means that there is such an asymmetry. But whatever asymmetry there is is due to the fact that most of the time we do not settle on a goal until we consider at least some of the problems which its realization may give rise to. So that agents already have a settled, though revisable view that this goal of theirs is worth pursuing, even though the means are not without costs and difficulties.²² Therefore, barring occasions where there is a case for revising that judgment, agents have an asymmetric attitude to the ends and the means, that is they do not think that they should abandon their ends because of the difficulties or costs of pursuing the means. The Schema has nothing to do with the asymmetry. It is a result of what constitutes a rational adoption of ends in most circumstances. If, or where, agents adopt ends without any attention to the means no asymmetry exists. Schroeder attributes other advantages to Schema Detach. They rely on the odd assumption that, as I did not explain the standards of rational functioning I alluded to, they cannot be explained except on the assumption that Detach is

²² This is so even if, as illustrated in Section Two, we have not yet adopted any instrumental intentions. Even in such cases, we usually have a rough idea of the problems various instrumental plans may involve.

valid.²³ He wrote more on these matters elsewhere, but this is not the place to discuss his views in detail.²⁴

Schroeder is right that my argument for the third thesis (that having goals is not and does not provide reasons) is indirect. That is the nature of most arguments for a negative existential. It seems natural to argue for the view that having goals does not provide reasons by showing that there is no reason to think that it does. That is what I tried to do. But I should have emphasised one point and set the argument against a more explicit background.

The point I should have emphasised is that from the point of view of an agent, the reasons to take actions which facilitate realization of his ends are the reasons for having the ends, and not his having the ends: I have reason to take medication which will cure my illness, and it is not that it is my goal to cure my illness, but that the illness threatens my health, interferes with my life, etc. What justifies my having the goal justifies pursuing the means. People, in other words, are not megalomaniac, or stubborn, insisting that their will must prevail (though sometimes they are). They set themselves to pursue goals because they take them to be worthwhile and therefore worth taking the means to achieve.

This point is, like all the others, not conclusive, and more than the others a matter of how you see things. Schroeder and others just do not see them that way. So let me turn to the considerations which explain how I see the argument: That one has an intention or a goal is just a non-evaluative fact like any other. How can one's intention or attitude, or an expression of such intention, or its communication, create reasons? This applies even when the intention is to create a reason (as in promising, or commanding). The fact that the tree leans northward is not a reason for me to go there, and how should my intention to go there be such a reason?²⁵ Of course there is a difference. The tree's orientation is unlikely to make me go north, whereas my

²³ Sobel, too, thinks that The Myth is an attempt to explain the normativity of those standards of rationality. It has much more modest ambitions than both of them assume. Schroeder rightly thinks that practical reasons are susceptible to an analogue of "undercutters" in epistemic reasons. Some 30 years ago I drew attention to the fact that some considerations – though not reasons themselves – are, what I then called, "weight-affecting" considerations. Other facts cancel reasons we had (e.g., being released from a promise). Schroeder's suggestion that nothing can cancel a reason once it is there – it can only diminish in weight or stringency – lacks plausibility. But in any case I fail to see how, were he right in that, the point could support the validity of Detach.

²⁴ Though I should say that I did follow his recommendation to examine whether desires are or provide reasons, and concluded that they do not. See *Engaging Reasons*, especially ch. 5. Schroeder, of course, disagrees.

²⁵ As Michael Bratman pointed out to me, I misled readers into thinking that the problem with thinking that people's goals provide reasons is that it means that they can manufacture reasons to perform immoral acts simply by intending to perform them. I meant these cases to be merely a vivid illustration of the underlying problem.

intention is likely to do that. But that is not a difference in normativity. It does not explain how my intention can be a reason to go north.

In general form, the answer is familiar: facts provide reasons when general evaluative or normative considerations determine that they do: that I am driving a car imposes on me a duty, which derives from the implications (to driving cars) of general considerations about our responsibilities toward others. Similarly our intentions, goals and their expression and communication impose duties when background considerations determine so. I have written extensively about how authorities can impose duties and confer rights just by their say so, and how promising can impose duties. I have also argued²⁶ that certain ways of forming intentions, broadly speaking by reaching decisions following deliberation of the merits, provide exclusionary reasons. Decisions, I argued, are reasons not to unsettle one's intentions, thus making refusal to reconsider them rational under some circumstances in which it would have been irrational had the intention been formed some other way. In each of these cases, the considerations which determine that such intentions or their expression provide reasons also set limits to the circumstances in which they do. And, in each of these cases, different considerations are at play. Section Four of *The Myth* considers further ways in which goals (rather than mere intentions) one has can affect one's reasons, and again the inquiry points to some considerations which show that in some ways they do. Far from having a blank opposition to the thought that goals, intentions or their expression and communication can provide reasons, I have done much to show how in many cases they do just that. Unfortunately there are no similar arguments which could justify a blanket conclusion that having goals or intentions provides reasons. That is where the negative argument, showing the failure of known attempts to establish that goals provide reasons, takes over.

Joseph Raz
Balliol College, Oxford University & Columbia University
joseph.raz@balliol.ox.ac.uk

²⁶ *Practical Reasons and Norms* (Oxford: OUP 2nd ed., 1999), though I would have argued the case somewhat differently, and would have somewhat modified the conclusion there expressed had I revisited the issue today.