3

The Ideal of Orthonomous Action, or the How and Why of Buck-Passing

MICHAEL SMITH

1. Two approaches

Imagine trying to explain to a group of people what a philosopher is. There are at least two approaches you could take. One would be to describe what an ideal philosopher does: the sorts of questions he thinks about, the methods he uses in answering them, the level of detail and precision he demands of his answers, how he deals with those who disagree with him, and so on. Having fixed the ideal, you could explain that non-ideal philosophers are approximations to the ideal. The other, much more flat-footed, approach would be to describe what all philosophers, ideal and non-ideal alike, have in common, and how they differ from non-philosophers.

Some might wonder whether there is any difference between these two approaches. If the only way to say what all and only philosophers have in common is to describe them as approximations—some more perfect, others less perfect—to the ideal, then this is indeed the case. However, if approximations to the ideal are such in virtue of their possession of some further (possibly disjunctive) lowest-common-denominator property then the two approaches are genuinely different. Moreover, if the two approaches are genuinely different, then, though both could be pursued in good conscience, a good question to ask is whether there are reasons to prefer one rather than the other, at least in certain circumstances.

The answer is that there may well be. Suppose, for example, that your task is to explain to some group of people what a philosopher is, but that you know in advance that you will disagree with them about some crucial feature of the ideal philosopher. Perhaps the ideal philosopher thinks about interesting problems, but their conception of which problems are of interest is very different to yours. In these circumstances, there would be at least one reason to take the lowest-common-denominator approach, as that approach wouldn't require you to take a stand on which problems are interesting. The lowest-common-denominator approach allows you to dodge this particularly thorny normative question.

This is not to say that the lowest-common-denominator approach allows you to dodge every normative question. For example, it might require you to specify how bad at doing philosophy one would have to be in order not to count as a philosopher at all. But since this is a question to which you would need an answer even if you took the idealization approach, the mere fact that the lowest-common-denominator approach embroils you in this normative controversy doesn't count against it. The earlier problem with the idealization approach thus wasn't that it makes you confront thorny normative issues, whereas the lowest-common-denominator approach doesn't; the problem was rather that it embroils you in all of the normative controversies in which the lowest-common-denominator approach embroils you plus more.

With this in mind, let's now ask a more familiar philosophical question, the question that sets the scene for this chapter. What is an action? According to the best-known version of the so-called 'standard story of action', a story we have inherited from David Hume via Donald Davidson, a subject's actions are those of his bodily movements that are done because he wants certain things and because he believes that he can achieve those things by moving his body in the ways he does (Hume 1777; Davidson 1963, 1971). According to a slightly less well known, but in my view much more plausible, version of the same story, which we owe to Carl G. Hempel (1961), a subject's actions are those of his bodily movements that are done because he wants certain things, because he believes that he can achieve those things by moving his body in the ways he does, and because he exercises his capacity to be instrumentally rational, so deriving an instrumental desire to move his body in the ways he does.

Any fully spelled out version of the standard story of action owes us an account of belief, desire, and the capacity to be instrumentally rational. The most plausible such accounts also derive from Hume. What makes a psychological state a belief, or a desire, or the capacity to be instrumentally rational, is its distinctive functional role. As Robert Stalnaker puts it, in perhaps the pithiest formulation:

> Belief and desire . . . are correlative dispositional states of a potentially rational agent. To desire that *P* is to be disposed to act in ways that would tend to bring it about that *P* in a world in which one's beliefs, whatever they are, were true. To believe that *P* is to be disposed to act in ways that would tend to satisfy one's desires, whatever they are, in a world in which *P* (together with one's other beliefs) were true. (1984: 15)

An agent who acts is thus one in whom these two behavioral dispositions interact in a manner that makes it appropriate to describe him as having exercised his capacity to be instrumentally rational (for more on this, see Smith 2003a). The concepts of belief, desire, rational capacity, and action are therefore all inter-defined. What marks them out are the distinctive ways they each relate to each other and the world.

Does the standard story of action take a lowest-common-denominator approach or an idealization approach to explaining what an action is? As I understand it, the standard story takes a lowest-common-denominator approach. In doing so, the standard story, much like the lowest-common-denominator approach to explaining what a philosopher is, requires us to take a stand on certain thorny normative and non-normative issues. For example, Hempel's version of the standard story commits us to the view that there is a norm of instrumental rationality; to the view that that norm governs relations between an agent's non-instrumental desires and means-end beliefs; to the view that, whenever there are actions, the agent in question possesses the capacity to be instrumentally rational to some degree; and to the view that, whenever there are actions, the agents of those actions exercise their capacity to be instrumentally rational, putting their non-instrumental desires and means-end beliefs together so as to derive instrumental desires to act in the ways in which they do. This last is what approximations to the ideal all have in common, or so the standard story tells us.

But though Hempel's version of the standard story of action takes a stand on these thorny normative issues, it remains non-committal on a whole host

of others. Since these are issues on which someone who takes the idealization approach must commit himself, this is the standard story's advantage. For example, someone who takes the idealization approach is committed to telling us which non-instrumental desires and means-end beliefs an agent acts on if he performs the ideal action—which dispositions he ought to have—something about which those who tell the standard story can remain totally silent. Indeed, those who tell the standard story needn't commit themselves to there being any norms at all governing an agent's non-instrumental desires, still less to specifying what the content of those norms is. The reverse is not true, however. Someone who takes the idealization approach, at least when he tries to identify what approximations to the ideal are, has no choice but to take a stand on the sorts of normative issues on which those who tell the standard story take a stand. When explaining what an action is to an audience with whom you have profound normative disagreements, it might therefore be more productive to tell the standard story.

Some might think that the standard story is plagued by problems of its own. Whereas the idealization approach signals that the border between actions and non-actions might be a vague matter—what counts as an approximation is, after all, often a vague matter—the standard story might seem committed to denying such indeterminacy. The whole idea behind the standard story, it might be thought, is to identify a bright line dividing actions from non-actions. But on a charitable reading, the standard story is not committed to there being a bright line. When an agent moves his body not because of anything he believes, but because he is in the belief-like state of vividly imagining something, then he should count as doing something that is at least a borderline case of action even by the lights of the standard story. The same is true when an agent moves his body because of some belief-mediated urge, rather than an ordinary desire. Borderline cases of action are thus predicted by the standard story, not counterexamples to it.

The upshot is that there are two approaches we might take to explaining what an action is and that these two approaches are also complementary. In earlier work, I have defended a version of the standard story. My reason for doing so has in part been the hope that we might get agreement on what an action is without confronting the difficult normative issues raised by the idealization approach. In this chapter, however, I wish to switch tack and

say a little about how we might take the alternative and complementary idealization approach instead. The idea I wish to explore is that actions, as identified by the standard story, are approximations to what Philip Pettit and I have elsewhere called *fully orthonomous* actions: that is, they are approximations to actions that are performed by agents who are under the rule of what's correct, as opposed to what's incorrect (Pettit and Smith 1996, Smith 2004). After briefly explaining this idea, I will zoom in and clarify some key aspects of orthonomous actions, so understood. Doing so will provide a background against which I can explain some misgivings I have concerning what Jonathan Dancy has to say about buck-passing conceptions of both rightness and values (2000).

2. Orthonomous agents are ruled by what's correct in which respect?

Fully orthonomous actions are those performed by agents who are themselves fully orthonomous, where agents are orthonomous to the extent that they are ruled by what's correct, as opposed to what's incorrect. An agent's being ruled by what's correct is thus a matter of his being sensitive, in his actions, to the way things are. He does what he does because of his appreciation of the way things are, not because of ignorance or error. This is why what he does counts as an idealized form of action.

An initial difficulty with stating this idea more precisely is that agents can be sensitive, in their actions, to the ways things are in many different respects. We must therefore be clear which respects we have in mind. Consider the following possibilities. There are agents who act in the light of their full sensitivity to all of the reasons that there are, both the moral reasons, if there are any such reasons, and the non-moral reasons too, that have any bearing at all on what they are to do; there are those who act in the light of their full exercise of their rational capacities, where this may or may not be a matter of their being fully sensitive to all of the reasons that there are; there are those who act in the light of a full sensitivity to specifically moral considerations, where this may or may not be a matter of their being fully sensitive to all of the reasons that there are and fully exercising their rational capacities; there are those who act in the light of a full sensitivity to

the legal code; others who act in the light of a full sensitivity to considerations of style; yet others who act in the light of a full sensitivity to considerations of humour; and so we could go on. When we describe a fully orthonomous action, and suppose that such an action is one that is performed by an agent who is ruled by what's correct, as opposed to what's incorrect, it follows that we have to make a choice about the relevant criteria of correctness. Correctness in which respect(s)?

One suggestion would be that we are supposed to describe those actions whose agents are ruled by what's correct in every respect. But this isn't a very promising suggestion, as it is not the case that, for every situation in which an ideal agent might find himself, there is an action that is correct in every respect. Humour and style, on the one hand, and morality, on the other, all too often pull in opposite directions from each other. Another possibility is that we are supposed to describe those actions whose agents are ruled by the exemplary mix of features that define what it is to be ruled by what's correct along all of the different dimensions. But this isn't a very promising suggestion either, as the idea of an exemplary mix makes dubious sense. Along which dimension is the mix supposed to be exemplary? Every dimension? Obviously not. But in that case, which? The agent who is moved by the mix that is exemplary as judged from a standpoint of sensitivity to all of the reasons that there are is presumably just the agent who is fully sensitive to all the reasons that there are; the agent who is moved by the mix that is morally exemplary is presumably just the agent who is fully sensitive to moral considerations; and so on.

It seems that we therefore have no alternative but to privilege some of these ideals over the others. In one sense, though, this makes the choice much easier. Since we cannot imagine individuals acting at all who are totally insensitive to reasons and who lack all rational capacities—people who act must display some degree of differential sensitivity to evidence that their environment is one way rather than another—we should suppose that ideal actions are, by their very nature, the sorts of things that are done by agents who are fully sensitive to all the reasons that there are and who fully exercise their rational capacities. Actions are not, in this same sense, defined by reference to the sorts of things done by those who are sensitive to considerations of style or humour, as we seem to have no problem at all imagining agents who have no sense of style or humour, and many insist that the same is true of agents who are insensitive to moral and legal

considerations. Of course, this doesn't show that ideal actions, or certain ideal actions, won't be stylish, or humorous, or morally or legally exemplary. But it does suggest that, if they are, then that will be because being stylish, or humorous, or morally or legally exemplary, is itself at bottom a matter of being sensitive to reasons and exercising rational capacities.

Here, then, is a preliminary attempt to describe what a fully orthonomous action is like:

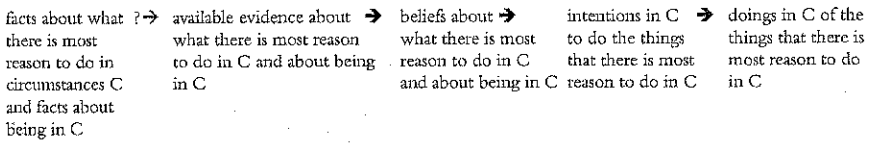


Figure 1 Fully orthonomous actions—1st try.

For agents to be ruled by what's correct, there must be facts about the various things that there is reason for them to do in their circumstances, and those facts must somehow combine to fix what there is most reason for them to do in their circumstances; there must be evidence of what those facts are when they are in those circumstances, and that evidence must be available to them; and they must exercise their capacities and actually access that evidence. How exactly we should conceive of the circumstances that are relevant to what agents have reason to do, and hence what's required for agents to access evidence that they are in those circumstances, will be a matter for further discussion presently. This is why, for the time being, the '?→' in figure 1 signals no commitment on that issue.

Having accessed this evidence, agents who are ruled by what's correct must respond to that evidence by forming judgements to the effect that the reasons that there are in their circumstances are those supported by the evidence—the '→' in figure 1 signals the exercise of rational capacities that's required for this transition between accessing evidence and judgements to be made—and they must go on to intend to do, and then to do, what they judge themselves to have most reason to do in their circumstances, once again by exercising their rational capacities. How we should conceive of reasons for action, and hence what's required for agents to access evidence that there are facts about reasons, will also be a matter for further discussion presently.

3. What are the circumstances in which fully orthonomous actions are performed?

If an agent is to act orthonomously, then non-normative facts about his circumstances—these are what's referred to in figure 1 as 'circumstances C'—must be available to him via evidence and he must access that evidence. But how exactly should we conceive of the circumstances that are relevant to what an agent has reason to do?

Consider a doctor with a patient who has a debilitating but non-fatal disease (compare Jackson 1991). Suppose that the best medical science tells him that either drug A or drug B will cure his patient completely, but that it doesn't tell him which will do the trick; imagine that it also tells him that whichever doesn't do the trick will kill him. Suppose further that the drug that will in fact cure his patient is drug A, and that there is a further drug, C, which though it won't cure his patient, is well known to ameliorate his disease's symptoms significantly. Given all of these non-normative facts about the doctor's circumstances, which drug does he have most reason to give? Cases like this force us to ask whether agents' circumstances, in the sense relevant to what they have reason to do, are fixed by the way the world is, or whether they are instead fixed by the way that they should believe the world is, given the available evidence.

If agents' circumstances are fixed by the way the world is, never mind what the available evidence suggests, then this suggests that the doctor has most reason to give drug A. If we think of agents' circumstances in this way, then there would always be a potential gap between agents' being in certain circumstances and its being available to them that they are in those circumstances. Even agents who respond perfectly rationally to the evidence available to them may fail to know what their circumstances are, as the available evidence might mislead. Their non-culpable ignorance would constitute some sort of excuse. Since a distinctive feature of orthonomous agents is that they know what their circumstances are, we would therefore need to signal the fact that they are not non-culpably ignorant in this way in figure 1. The '?→' would have to be replaced with a 'K→' which would tell us that the world has conspired to make the facts about the orthonomous agent's circumstances available to him in a manner suitable for knowledge.

If, on the other hand, agents' circumstances are fixed by the available evidence, then that would suggest that the doctor has most reason to give

drug C. Though the available evidence might mislead agents about how things actually are in the world in which they act, it could not mislead them about their circumstances in the sense relevant to what they have reason to do, as there would be no gap between the circumstances in which agents find themselves, in this sense, and the availability of evidence that they are in those circumstances. Agents who respond perfectly rationally to the available evidence, as orthonomous agents do, could not fail to know what their circumstances are; there could be no non-culpable ignorance of circumstances. We would need to signal this fact in figure 1 by replacing the '?→' with '=' which would represent the constitutive relationship between agents' circumstances, in the relevant sense, and the evidence available to them about their circumstances.

Which of these is the correct way to think about agents' circumstances? Though I am not absolutely certain, my inclination is to think that agents' circumstances are fixed by the way the world is, not by the way that the available evidence suggests that it is. My reason for thinking this is that I cannot see how we could justify taking an asymmetrical attitude towards the non-normative circumstances in which agents act, on the one hand, and the facts about what there is reason for them to do in those circumstances, on the other (see also Smith 2009a). If we should suppose that one of these is epistemically constrained, then we should suppose that the other is epistemically constrained too. But if the world itself has a normative nature, then it seems plain that it is, *inter alia*, that normative nature that fixes what agents have reason to do, not the possibly misleading evidence agents might have about that normative nature. Similarly, it seems to me that we must suppose that the non-normative circumstances in which agents find themselves are fixed, *inter alia*, by the non-normative nature of the world, not by the possibly misleading evidence agents might have about the world's non-normative nature.

So, at any rate, I am inclined to think. But since nothing in what follows will turn on this, I will say no more about it. The rest of the picture of orthonomous action I go on to sketch could easily be reconceived in terms of the alternative conception of agents' circumstances if it turns out that I am wrong and an asymmetrical attitude can be justified.

4. Reasons for action or reasons for intention?

If an agent is to act orthonomously then, however we should conceive of the circumstances in which he finds himself, there must also be facts about what there is most reason to do in various circumstances, including the circumstances in which the agent finds himself, and these facts about reasons for action must also be available to him via evidence that he accesses. Given that figure 1 is supposed to be portraying the structure of orthonomous action, it therefore represents facts about what there is reason to do in various circumstances as themselves the basic normative facts to which agents respond.

In order to test the plausibility of this, consider a case in which we might suppose that there are reasons for agents to act in certain ways, and let's ask whether it is plausible to suppose that facts about such reasons are basic. People like me who have children have a reason to make sure that their children are safe and well, or so I will assume. But is it a basic normative fact that people like me have such reasons, or is this reason for action explained by something more basic? Thomas M. Scanlon provides what seems to me to be a compelling reason for supposing that such reasons are not basic. He points out that among the mental states that people possess, there is a distinctive class of what he calls *judgement-sensitive attitudes*. Judgment-sensitive attitudes are those

> that an ideally rational person would come to have whenever that person judged there to be sufficient reasons for them, and that would, in an ideally rational person, 'extinguish' when that person judged them not to be supported by reasons of the appropriate kind. (Scanlon 1998: 20)

The reasons to which judgment-sensitive attitudes are sensitive are what he calls reasons in the 'standard normative sense', the paradigmatic examples of which are considerations that support the truth of our beliefs, but other examples include intention, desire, hope, fear, admiration, respect, contempt, indignation, and so on (pp. 20–1).

What is important about the various judgment-sensitive attitudes, according to Scanlon, is that they 'constitute the class of things for which reasons in the standard normative sense can be asked or offered' (p. 20). Among mental states, the judgement-sensitive attitudes thus contrast with states like being in pain, or feeling dizzy, for which reasons in the standard normative sense cannot be given. (It is worth comparing judgement-sensitive

attitudes, as Scanlon characterizes them, with what Judith Jarvis Thomson calls the distinctive class of mental states with 'correctness conditions' (Thomson 2008: 132).) But judgement-sensitive attitudes also contrast with actions, and this may seem more problematic. For though actions are not attitudes, and so are not judgement-sensitive attitudes, they are the sorts of things for which reasons in the standard normative sense can be given. So do they constitute a counterexample to Scanlon's suggestion that judgement-sensitive attitudes constitute the class of things for which reasons in the standard normative sense can be given?

Scanlon (1998: 21) thinks not. As he puts it:

> Actions are the kinds of things for which normative reasons can be given only insofar as they are intentional, that is, are the expression of judgement-sensitive attitudes. Against this, it might be pointed out that (at least in normal cases) in order to intend to do something I must take myself to have a reason for *doing* that thing. So it might seem that reasons for action are, after all, primary, and reasons for intending are dependent upon them. But there is no real disagreement here. A reason for doing something is almost always a reason for doing it intentionally, so 'reason for action' is not to be contrasted with 'reason for intending'. The connection to action, which is essential to intentions, determines the kinds of reasons that are appropriate for them, but it is the connection with judgment-sensitive attitudes that makes events actions, and hence the kind of things for which reasons can sensibly be asked for and offered at all.

As Scanlon sees things, reasons for actions are thus nothing over and above reasons for intentions (I will say something about the exceptions Scanlon mentions in passing in a moment). There are therefore reasons for action in the standard normative sense only because there are reasons for intentions.

Assuming that Scanlon is right about this, as I think he is, it follows that we should revise figure 1. Given that our aim is to describe the structure of orthonomous action, we should replace all mention of *reasons for action* with talk of *reasons for intention*:

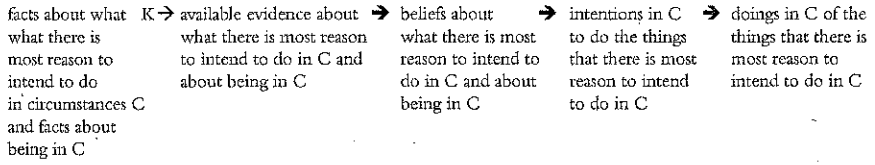


Figure 2 Fully orthonomous agency—2nd try.

Note that facts about reasons for actions are not explicitly mentioned in figure 2 at all. In order to interpret figure 2 as a representation of the way in which fully orthonomous agents respond to facts about reasons for actions, we have bring to it our knowledge that reasons for action decompose into reasons for intention.

5. Reasons for intention or reasons for desire?

But is it any more plausible to suppose that reasons for intention are basic normative facts, as suggested by figure 2, than it is to suppose that reasons for action are basic normative facts? In order to test this idea, consider once again our example.

If people have a reason to make sure that their children are safe and well, and if reasons for action decompose into reasons for intentions, then it follows that people have a reason to intend to make sure that their children are safe and well. Let's now imagine a situation in which my children are already safe and well, and robustly so, but let's suppose further that their being safe and well had, and has, nothing whatsoever to do with me. Their being safe and well is like manna from heaven. What should my reaction be? Would there be something for me to regret about this situation? The answer, I think, is that there would be nothing to regret. I should be delighted that my children are robustly safe and well, as their being safe and well is all that matters. It provides all of the relevant reasons that there are in the circumstances. Insofar as there is a reason for the intention to make my children safe and well, that reason thus seems itself to be explained by the reason for the desire I have that they be safe and well, a desire whose satisfaction brings with it, or perhaps constitutes, my being delighted.

How exactly would that explanation go? The explanation would go something like this: the world could be such that my children are safe and well; I have a reason to non-instrumentally desire that the world is such that my children are safe and well; given that one of the ways in which the world could be such that my children are safe and well is by my taking certain steps to make it that way, I have a reason to instrumentally desire that I take those steps to make it that way; and given that I can take those steps, I have a reason to intend to do so. Facts about reasons for non-instrumentally

desiring the world to be a certain way are in this respect seen to be more basic normative facts than facts about reasons for intentions. Indeed, we might well wonder whether reasons for intentions, any more than reasons for action, need to be mentioned explicitly in our description of fully orthonomous action at all. For there seems to be no real difference between reasons for instrumentally desiring to take certain steps, where those steps are ones that can be taken, and reasons for intending to take those steps.

It might be thought that this line of reasoning leads to a limited conclusion at best. Notwithstanding the fact that there would be much to be glad about, it might be thought that there would be something to regret if my children's being safe and well was like manna from heaven. After all, if my children's being safe and well was like manna from heaven, then it would not be possible for me to express my concern for them by making sure that they are safe and well. But, the thought might be, there plainly are reasons for me to engage in such expressive acts, reasons that are quite independent of the fact that one consequence of such acts is that my children are safe and well. Facts about reasons to intend to perform these expressive acts should themselves therefore be thought of as basic normative facts, facts that are not reducible to facts about reasons to non-instrumentally desire the world to be a way that it could be without my intending to do anything.

However this line of thought isn't really convincing. For one thing, though we do often say that we have reasons to perform expressive acts, it isn't entirely clear that the fact that we say such things should be taken at face value. Certain expressive acts, like acts of spontaneous affection for example, don't seem to be acts that we could succeed in performing at all if we did them with the intention of expressing spontaneous affection. These are presumably some of the exceptions Scanlon had in mind when he said that 'a reason for doing something is *almost always* a reason for doing it intentionally' (my emphasis). The upshot is that if we have reasons to perform acts of spontaneous affection, then these are not reasons to perform such acts intentionally. So should we suppose that they are reasons to perform such acts *unintentionally*? What on earth would a reason like that be?

When we say that there are reasons to perform acts of spontaneous affection, what we really have in mind, or so it seems to me, is one of two quite different things. One possibility is that we are saying that there are reasons to perform acts of spontaneous affection under some other description. There is a reason to, say, buy a gift, or plant a kiss, or whatever else it is

that we do intentionally when we act spontaneously. But if this is what we have in mind, then it is clear that the example isn't probative, as the reasons to buy a gift, or plant a kiss, or whatever, could well be reasons to intend to buy a gift, or plant a kiss, or whatever. But the other thing that we might really have in mind when we say that there are reasons to perform acts of spontaneous affection is that there are reasons to *desire* that we perform such acts under that very description, and this is probative.

What's peculiar about reasons to desire that we do things like perform acts of spontaneous affection, unlike reasons to desire things like my children's being safe and well, is that there is no way for our sensitivity to such reasons to lead us to acquire the desire for which they are reasons and then to act so as to satisfy the desire. Such reasons thus cannot be reasons for corresponding intentions. In order to perform acts of spontaneous affection I have to *act spontaneously*, not in the light of my appreciation of the reasons for desiring that I so act. Reasons for desiring that I perform acts of spontaneous affection cannot be converted into reasons to intend to act spontaneously because, in the circumstances in which I appreciate those reasons, my very appreciation of those reasons ensures that there are no ways in which the desire could lead me to act so as to satisfy the desire for which they are reasons.

Reasons for other expressive acts are different. Consider the reason I have to express my concern for my children. This is presumably one and the same as the reasons I have to want the world to be one in which my children know that I love them and would be willing to make all sorts of sacrifices to ensure that they are safe and well. To be sure, in the world as it actually is, my children's main source of evidence about such things lies in the actions I perform. Moreover, I could perform those acts fully cognizant of the reasons for so desiring. It therefore seems that, in the world as it actually is, I do have reasons to intend to perform those acts in order to convey that knowledge. But there is nothing necessary about any of this. It is logically possible for there to be a world in which my children know how much I love them, and the sacrifices that I would make for them, without my doing anything to convey that knowledge to them, and indeed without anyone else's doing anything to make that happen either. It seems that we should therefore suppose that the reasons I have to intend to perform such expressive acts are explained by the reasons I have for wanting the world to be a certain way, where this is a way that the world could be without my

acting at all. The reasons to intend to perform such expressive acts are not themselves basic.

This is in turn significant. For it suggests is that, quite generally, reasons for action are explained by reasons to non-instrumentally desire the world to be a certain way, where it is left entirely open whether it could in principle get to be that way by agents' being sensitive to those reasons and, as a result, by their taking steps to make it that way. That will depend entirely on the way that there is reason to want it to be. In many cases, the world could be the way that there is reason to want it to be without those who have such reasons being sensitive to those reasons and, as a result, doing something to make it that way. But in other cases this may not be possible. Perhaps there are some ways that we have reason to non-instrumentally desire the world to be where it could only get to be that way if those who have such reasons were sensitive to those reasons and, as a result, do something to make it that way.

Once we acknowledge that this is so, it follows that our conception of orthonomous action must sharply separate out the role played by facts about the steps that an agent could take in his circumstances—ultimately, the movements he makes with his body—to make it the way that he has reason to non-instrumentally desire it to be. When the world is already the way that there is reason to non-instrumentally desire it to be, there won't be any facts about what the agent could do in his circumstances to make it that way. This is a corollary of what we learned by thinking about the reasons that there are for desiring that my children be safe and well when their being safe and well is like manna from heaven. And even when there are such facts, these facts may not be able to play the role that they have to play in action. This is a corollary of what we learned by thinking about the reasons for desiring that we perform acts of spontaneous affection.

When an agent acts orthonomously, he therefore has to be sensitive both to facts about the ways there is reason to non-instrumentally desire the world to be and to facts about the steps that he could take in his circumstances to make it these ways. He must acquire corresponding non-instrumental desires and beliefs; he must put his resultant non-instrumental desires and beliefs together to form instrumental desires; and he must act on those instrumental desires. We must therefore abandon the conception of orthonomous action represented in figure 2 and suppose that orthonomous actions are more properly represented by a two-track process as follows:

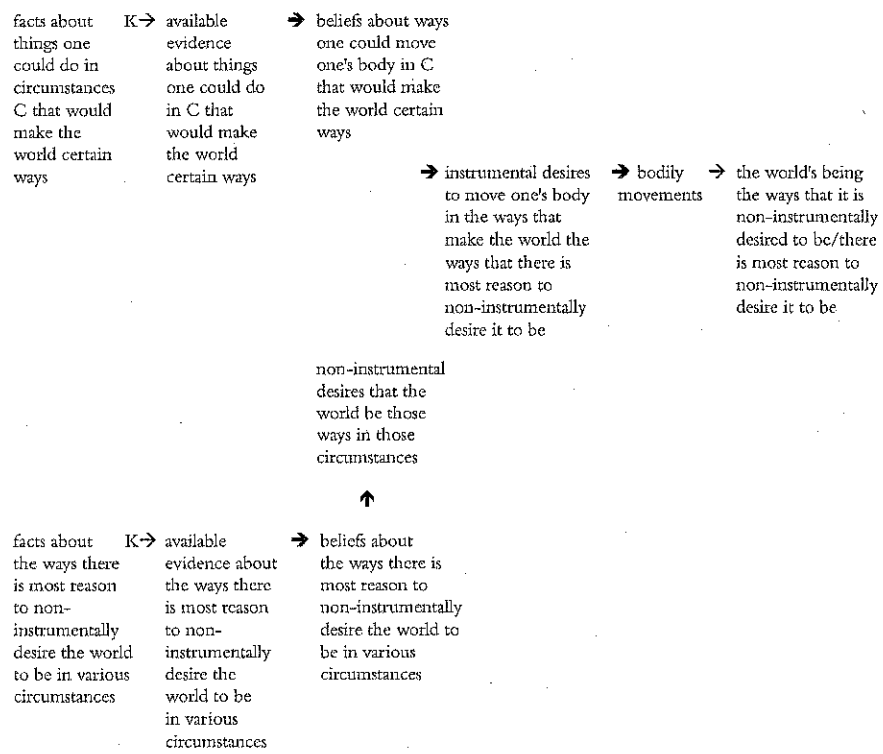


Figure 3 Fully orthonomous action—3rd try.

Though neither reasons for actions nor reasons for intentions are mentioned explicitly in this representation of orthonomous action, both are implicit. We see that figure 3 represents an orthonomous agent's sensitivity to reasons for intentions when we remember that reasons for intentions are nothing over and above reasons for instrumentally desiring, where the instrumental desires in question are desires about steps that can be taken by the orthonomous agent in his circumstances to satisfy the relevant non-instrumental desire for which there are reasons. And it represents an orthonomous agent's sensitivity to reasons for action because, as Scanlon points out, reasons for action are nothing over and above reasons for intention.

Let me sum up. Though we might initially have thought that orthonomous agents are at bottom sensitive to reasons for performing certain actions, we have seen that this is not so. The reasons for action to which orthonomous agents are sensitive decompose into reasons for intention, and

these reasons in turn decompose into reasons to desire that the world be certain ways and into reasons to believe that there are things that orthonomous agents can do in their circumstances to make the world those ways. Though these reasons may decompose further, this is decomposition enough for us to see the attractions of the standard account of action.

The standard account, remember, is supposed to be an account of the feature shared by non-ideal actions in virtue of which they are approximations to the ideal. According to figure 3, actions are non-ideal for many different reasons. They might be non-ideal because the facts to which orthonomous agents are sensitive are not available to non-orthonomous agents. Or they might be non-ideal because, though these facts are available to them, non-orthonomous agents do not form their beliefs in the light of the available evidence. Or they might be non-ideal because, though non-orthonomous agents form their beliefs in the light of that evidence, they do not form their non-instrumental desires in the light of these beliefs, and so the non-instrumental desires on which they act are not those for which they believe that there is most reason. And, of course, actions might also be non-ideal because of every possible combination of these various forms of dysfunction.

So do non-ideal actions all share a feature in virtue of which they are approximations to ideal actions? They do indeed. Look again at the various nodes in figure 3 where we might find dysfunction. Non-ideal actions all share the feature of being the product of what their agents non-instrumentally desire to be the case and what they believe they can do to satisfy their non-instrumental desires in their circumstances. This is just the standard account of action. Note, however, that we have arrived at this account not by identifying its requirements as the lowest common denominator shared by ideal and non-ideal actions alike, but rather by independently describing ideal actions and then identifying its requirements as the feature shared by approximations to the ideal. To my mind, this underscores the plausibility of the standard account.

6. Dancy on buck-passing

With this account of ideal actions before us, we are in a position to describe some problems for what Jonathan Dancy has to say about buck-passing

views of goodness and rightness. It might initially seem that these two topics have little to do with each other, but, as we shall see, this is far from being so.

Dancy writes:

> To claim that something is good is to claim, according to the buck-passing view, that it has features that give us (*pro tanto*) reasons to take certain attitudes to it. Similarly, we might say, to claim that some action is right is to claim that it has features that give us overall reasons to do it. This is a buck-passing view because it holds that the normative force of a claim that something is right or wrong is inherited from that of the reasons which it asserts to be present. (2000: 166–7)

Though Dancy thinks that a consistent buck-passer would pass both bucks—he would pass the goodness buck to reasons for *attitudes* and the rightness buck to reasons for *actions*—he also insists that there is good reason not to be a consistent buck-passer. An action's being right is a function of the reasons in favour of doing it, according to Dancy, but something's being good is not just a matter of the reasons in favour of taking certain attitudes towards it.

Let me begin with a quibble. Dancy's official argument for, and statement of, the buck-passing view about rightness turns on his suggestion that judgements of rightness are verdictive in the following sense.

> In deciding whether an action is right, we are trying to determine how the balance of reasons lies. Our conclusion may be that there is more reason (or more reason of a certain sort, perhaps) to do it than not to do it, and we express this by saying that it is therefore the right thing to do. The rightness-judgement is verdictive; it expresses our verdict on the question how the reasons lie. It is incoherent, in this light, to suppose that the rightness can add to the reasons on which judgement is passed, thus, as one might say, increasing the sense in which, or the degree to which, it is true. And the same is true of wrongness. (Dancy 2000: 166)

When Dancy says that the same is true of wrongness, I take it he means that wrongness is verdictive in the very same sense. In deciding whether an action is wrong, we are trying to determine whether there is more reason (or more reason of a certain sort) not to do it than to do it, and we express this by saying that it is wrong.

But if an action's not being right entails that it is wrong, and if an action's not being wrong entails that it is right, then Dancy's statement of the buck-passing

view of rightness is misleading. When we say that an act is right, we would have to be expressing the negation of the thought we express when we say an act is wrong (compare Wedgwood 2008). In other words, our thought would have to be that it is not the case that there is more reason not to do the act in question than to do it. (Given Dancy's parenthetical remark, it might be better to say that our thought would have to be that it is not the case that the action is of a certain kind, and, among acts of that kind, it is not the case that there is more reason not to do it than to do it. From here on, I will ignore this complication.)

What's important about this way of formulating the buck-passing view of rightness is that it allows that there may be no such thing as *the* right thing to do in a situation—note that this is what Dancy is in fact analysing in the passage quoted—as a number of different acts could equally be right. If, say, I am obliged to ensure that my children are safe and well, and I could do this either by taking out an insurance policy or by investing in stocks and bonds, then there is more than one right thing to do. I act rightly whether I take out the insurance policy or invest in stocks and bonds because it isn't the case that there is more reason not to perform these acts than to do so. But neither act is such that there is more reason to do it than not to do it, so neither satisfies Dancy's criterion of rightness.

With this quibble on the table, I will ignore it in what follows. For if the buck-passing view of rightness just described is correct, then though an act's being right isn't a matter of there being more reason to do it than not to do it, an act's being *the* right act—that is, the one and only right act—will indeed be a matter of there being more reason to do it than not to do it. So let's now focus on that suggestion and Dancy's arguments for it and the correlative view of wrongness. Dancy thinks that an act's being the right act is a matter of the balance of reasons favouring doing it. In other words, remembering now the initial discussion of figure 1, it is a matter of there being most reason to so act. But as the subsequent discussion and revisions of figure 1 make clear, once we pass the rightness buck to what there is most reason to do, we are led ineluctably to pass it much further. The rightness buck passes next to reasons for intention, as in figure 2, and then to reasons for non-instrumental desires and means-end beliefs as in figure 3.

This casts Dancy's initial discussion of consistent buck-passing in a very different light. The consistent buck-passer wouldn't just pass the rightness buck to reasons for action and the goodness buck to reasons for attitudes,

he would pass both the goodness buck and the rightness buck all the way to reasons for attitudes, and, moreover, to reasons for the very same attitudes. Being good and being right would both be a matter of there being reasons for non-instrumental desiring. This is somewhat obscured in Dancy's discussion, as he follows Ewing in emphasizing, against Ross, that if we suppose that goodness is the property of having some feature that provides reasons for attitudes, then we should think that that set of attitudes includes all of

> choice, desire, liking, pursuit, approval, admiration . . . When something is intrinsically good, it is (other things being equal) something that on its own account we ought to welcome, rejoice in if it exists, seek to produce if it does not exist. We ought to approve its attainment, count its loss a depravation, hope for and not dread its coming if this is likely, avoid what hinders its production, etc.
>
> (Ewing 1947: 149)

But while all of this is true, and while Dancy is surely right that the diversity of attitudes undermines Ross's reasons for not being a buck-passer about goodness, we have also seen that there is a specific attitude that is in play when acts are right, namely, the attitude of intrinsic desiring.

The upshot is that since rightness reduces to reasons for action, and since reasons for action reduce to reasons for non-instrumental desiring, and since this is *inter alia* what goodness consists in, so rightness reduces to goodness. Consistent buck-passing thus entails that the debate between consequentialism and deontology is resolved decisively in favour of consequentialism, a point that Dancy himself notes and to which we will return below. Those attracted to deontology should therefore give up their opposition to reducing rightness to goodness and insist, instead, that there are more goods than those allowed by typical consequentialists. In particular, they should insist that there are goods that underwrite the various deontological duties (Sen 1982, 1988; Broome 1991; Dreier 1998; Smith 2003b, 2009).

However, as noted earlier, Dancy doesn't think that we should be consistent buck-passers. As I understand it, Dancy is opposed to consistent buck-passing because he thinks that there is a more plausible view about goodness in the offing. Instead of supposing that goodness is the property of having some feature that provides us with a reason to have some pro-attitude, he thinks that goodness may be a distinct property, albeit one

which is grounded in the very same feature that provides us with a reason to have some pro-attitude. As he puts it, the alternative view that needs to be ruled out before we embrace consistent buck-passing is that 'reasons and values are distinct but may have the same grounds' (2000: 165).

It might be thought that this view could be ruled out on grounds of parsimony—or, to use Dancy's preferred term, on grounds of 'theoretical neatness' (p. 165)—as it allows that goodness and the property of having some feature that provides us with a reason to have some pro-attitude are necessarily co-extensive. But I take it that, at least as Dancy sees things, the two properties are not co-extensive.

> The real point, I think, is one about the polyadicity of rightness and of goodness—and this is a point that translates into thoughts about the polyadicity of reasons. Let us allow, without asking why for the moment, that rightness is a many-place relation. The point will then be that even if goodness is also a many-place relation, it has fewer places than rightness does and fewer than reasons do. Now if this is true, it cannot be correct to define goodness as the presence of reason-giving features. For the presence of reason-giving features will have more places in it, so to speak, than the goodness has.
>
> The reason for supposing that goodness is less polyadic than reasons is that reasons belong to, and are for, individuals. There are no reasons hanging around waiting for someone to have them. If the situation generates a reason for action, it must allot that reason to someone . . . But goodness is not like this. Something can be good or bad without specification of an agent. The desolation or destitution of someone is bad even if there is nobody around to do anything about it, nobody who has any opportunity to do anything about it, and so nobody who can be said to have a reason to do something about it. Someone's destitution, then, has features that would ground reasons for any agent suitably situated, but it does not follow that those features already ground reasons. And if it does not, we can be sure that to have value or disvalue is not itself to have reason-giving features.
>
> (pp. 170–1)

There may therefore be goodness and badness in circumstances in which the feature that would have grounded reasons, had circumstances been slightly different, doesn't ground any reasons. So, at any rate, Dancy argues. How should we respond?

The first response is that Dancy seems to misunderstand what buck-passing about goodness requires. According to the buck-passer, someone's destitution is bad just in case it has features that provide anyone with a reason to *be averse* to someone's being in a condition like that. And destitution may indeed have such a feature, despite what he says, for what it is like to be destitute may provide anyone with a reason to be averse to someone's being in a state like that, whether or not anyone can do anything about it. Of course, for this feature to provide people with a reason *for action*, there would have to be something that they could do about it in their circumstances. This was why figure 3 depicts orthonomous action as a two-track process. On one track there are reasons for the orthonomous agent to non-instrumentally desire that things be a certain way, where it is left open whether or not there is anything that can be done to make them that way, and on the other track there are reasons for the orthonomous agent to believe that there is something that he can do to make things that way in the circumstances in which he finds himself. Dancy is therefore right that someone's being destitute may be bad, even though no one has a reason to do anything about it, but he is wrong that this counts against the truth of being a buck-passer about the badness of his destitution.

A second response is required, however. For it might be thought that Dancy's real objection has nothing to do with supposing that the reasons in question are reasons for action, rather than reasons for aversion. His real objection is that the buck-passer is committed to there being reasons 'hanging around waiting for someone to have them'. If a situation generates a reason for aversion, then it must allot that reason to someone. But in the imagined situation, Dancy might say, this condition isn't satisfied. For no one is around to be allotted the reasons for aversion to destitution that the buck-passer postulates.

But is this right? I don't think so. According to the buck-passer, *everyone* has a reason to be averse to anyone's being destitute. Of course, some people may not know that a particular destitute person is destitute, and so they may not know that they have a reason to be averse to that particular person's being in the state that he is in. But this is hardly an objection to the buck-passing view, as the buck-passing view doesn't require that people have such knowledge of the destitution of particular people. And though it is presumably a priori knowable that everyone has a reason to be averse to anyone's being destitute, the buck-passer can also allow that some people lack this knowledge too. For even though a priori knowable, the buck-passer isn't committed to the empirical thesis that everyone does in fact know this particular a priori knowable truth. He can therefore allow that everyone has

a reason to be averse to anyone's being destitute even though some people do not know that this is so. Dancy's objection to buck-passing on the basis of the different polyadicity of reasons and values therefore fails to convince. We have been given no reason to suppose that reasons and values do differ in their polyadicity.

This connects with another of Dancy's objections:

> An action can be one's duty even though doing it has no value and its being done generates nothing of value. Standard examples here are of trivial duties. Suppose that I promise my children that I will tie my right shoelaces before my left shoelaces on alternative days of the week if they will do their homework without fuss. One can imagine arguing that though I ought to tie my right shoelaces before my left shoelaces today, since I did the opposite yesterday, my doing so has no value of any form. The buck-passing view rules this out in advance. To have value is to have reason-giving features, we are told, and since this is an identity statement it goes both ways. So to have reason-giving features is to be of value. So the deontological view expressed above is ruled out in advance of any significant debate. (Dancy 2000: 168)

But while this is so, as we have already noted, the real question is whether it's being so constitutes an objection to buck-passing, or merely a surprising consequence of it.

Let's begin by fixing on what it is about keeping even trivial promises, like the one Dancy imagines, that provides everyone with a reason to keep even them. According to one plausible view, which I will simply assume to be correct in what follows, even trivial promises create reasonable expectations, so, since what everyone has a reason to desire is that the reasonable expectations they create are met rather than unmet, it follows that even trivial promises create reasons (compare Scanlon 1998, Smith 2011). Of course, if this is right, then it follows that the value of a kept trivial promise is agent-relative, rather than agent-neutral. The value is agent-relative because, in specifying what it is that everyone has reason to desire, ineliminable reference is made to the agent himself: everyone has a reason to desire that the reasonable expectations that *he himself* creates are met rather than unmet. This means that there is a contrast with the disvalue of destitution, as just discussed. For in characterizing what it is that everyone has reason to be averse to, as regards destitution, no reference at all needed to be made to those who stand in a certain relation to that destitution. Everyone has a reason to be averse to destitution no matter what relation they stand in to that destitution. The disvalue of destitution is therefore agent-neutral.

When Dancy says that an 'action can be one's duty even though doing it has no value and its being done generates nothing of value', we should therefore reply that this assumes that the only value is agent-neutral value. Keeping a trivial promise can indeed be one's duty, even though its being done generates nothing of agent-neutral value. But since there is a reason to keep a trivial promise, it follows that keeping it does produce something of agent-relative value, in particular, the meeting of the reasonable expectation that the person who made that trivial promise created.

Moreover, now going on the offensive, it might further be replied that, absent the creation of something of value, whether agent-neutral or agent-relative—that is, absent the existence of a reason to desire or be averse to something or other, whether the thing in question has to be characterized with or without reference to an agent who bears some sort of special relation to that thing—there can be no obligation, because obligations entail reasons for action, and reasons for action reduce to reasons for intention, which in turn reduce to reasons for desire and aversion. The deontological view Dancy imagines is therefore best recast as a view about the existence of obligations grounded in agent-relative values, not as a view about the existence of obligations that aren't grounded in values.

Dancy in effect notes that a reply of this kind is available, but he rejects it:

> There is, however, a possible way out of this difficulty. It involves the introduction of agent-relative value. . . . Now I don't want here to go into the details of how this might be done, if indeed it can be done at all. My point at this juncture is going to be merely that, if we are to try to prevent the adoption of the buck-passing view from undermining a significant aspect of deontology by introducing a conception of agent-relative value, this is a considerable theoretical cost, and it is a cost, once again, that we have committed ourselves to paying just for the sake of theoretical neatness. What is more, many people doubt the coherence of the notion of agent-relative value in the first place. If the buck-passing view can only be sustained by introducing a piece of dubious philosophy, it is looking much less attractive. (Dancy 2000: 169–70)

The trouble with this response, however, is that it gets things the wrong way around.

Even Dancy can and should agree that there may be reasons for all of us to be averse to anyone's being destitute, no matter what relation we stand in to that destitution. And even he can and should agree that there may be reasons for all of us to desire to meet not just any old reasonable expectations, no matter what relation we stand in to those expectations, but specifically the reasonable expectations that we ourselves create. There is no incoherence or theoretical cost involved in supposing that reasons of each of these kinds exist. All that the buck-passing view of value does is parlay this view about reasons for attitudes into a view about values. It tells us that agent-neutral values consist in reasons of the former kind and that agent-relative values consist in reasons of the latter kind. The buck-passing view thus shows not just why there is no incoherence in the concept of agent-relative value, but also why there is no theoretical cost involved in supposing that such value exists alongside agent-neutral value.

References

Broome, J. (1991) *Weighing Goods* (Oxford: Blackwell).

Dancy, J. (2000) 'Should We Pass the Buck?' in A. O'Hear (ed.), *Philosophy, the Good, the True, and the Beautiful*, Royal Institute of Philosophy Supplement, 47: 159–74.

Davidson, D. (1963) 'Actions, Reasons, and Causes', in his *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), 3–20.

——(1971) 'Agency', in his *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), 43–62.

Dreier, J. (1993) 'Structures of Normative Theories', *The Monist*, 76: 22–40.

Ewing, A. C. (1947) *The Definition of Good* (London: Macmillan).

Jackson, F. (1991) 'Decision Theoretic Consequentialism and the Nearest and Dearest Objection', *Ethics*, 101: 461–82.

Hempel, C. (1961) 'Rational Action', in N. Care and C. Landesman (eds.), *Readings in the Theory of Action* (Bloomington, IN: Indiana University Press, 1968), 285–6.

Hume, D. (1777) *Enquiries Concerning Human Understanding and Concerning the Principles of Morals* (Oxford: Clarendon Press, 1975).

Pettit, P. and Smith, M. (1996) 'Freedom in Belief and Desire', *Journal of Philosophy*, 93: 429–49.

Scanlon, T. M. (1998) *What We Owe to Each Other* (Cambridge, MA: Harvard University Press).

Sen, A. (1982) 'Rights and Agency', *Philosophy and Public Affairs*, 11: 3–39.

——(1988) 'Evaluator Relativity and Consequential Evaluation', *Philosophy and Public Affairs*, 12: 113–32.

Smith, M. (2003a) 'Rational Capacities', in S. Stroud and C. Tappolet (eds.), *Weakness of Will and Varieties of Practical Irrationality* (Oxford: Oxford University Press), 17–38.

——(2003b) 'Neutral and Relative Value after Moore' *Ethics*, 113 (Centenary Symposium on G. E. Moore's *Principia Ethica*): 576–98.

——(2004) 'The Structure of Orthonomy', in J. Hyman and H. Steward (eds.), *Action and Agency*, Royal Institute of Philosophy Supplement 55: 165–93.

——(2009a) 'Consequentialism and the Nearest and Dearest Objection', in I. Ravenscroft (ed.), *Minds, Ethics, and Conditionals: Themes from the Philosophy of Frank Jackson* (Oxford: Oxford University Press), 237–66.

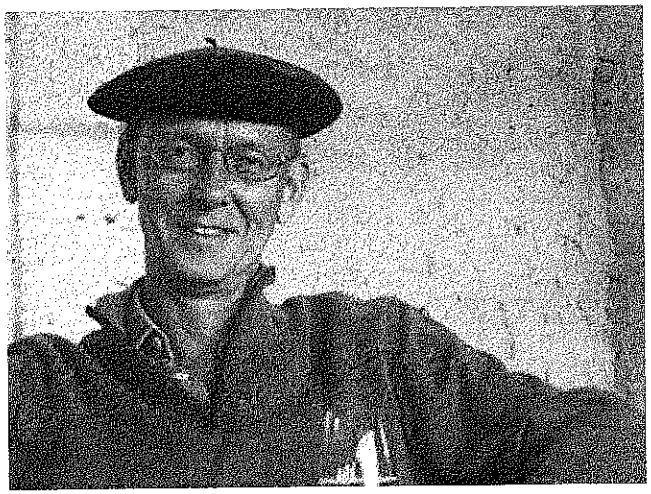
——(2009b) 'Two Kinds of Consequentialism', *Philosophical Issues*, 19 (Metaethics): 257–72.

——(2011) 'The Value of Making and Keeping Promises', in H. Sheinman (ed.), *Promises and Agreements: Philosophical Essays* (New York: Oxford University Press), 198–216.

Stalnaker, R. (1984) *Inquiry* (Cambridge, MA: MIT Press).

Thomson, J. (2008) *Normativity* (Chicago: Open Court).

Wedgwood, R. (2008) 'A Puzzle about "Right" and "Wrong"', *PEA Soup*, 11 May, <http://peasoup.typepad.com/peasoup/2008/05/a-puzzle-about.html>. Last accessed November 2012.



Jonathan Dancy. Photograph by Sarah Dancy.

Thinking about Reasons

Themes from the Philosophy of Jonathan Dancy

EDITED BY

David Bakhurst, Brad Hooker, and Margaret Olivia Little

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2013

Impression: 1

British Library Cataloguing in Publication Data

Data available

ISBN 978–0–19–960467–8

Printed and bound in Great Britain by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Contents