



Reasons, Reasoning, Reasonableness (2002)

by John Gardner
Professor of Jurisprudence
University of Oxford

& Timothy Macklem
Professor of Jurisprudence
King's College London

<http://users.ox.ac.uk/~lawf0081> <http://bit.ly/1t53Jxr>

This is an author eprint, which may not incorporate final edits.
The definitive version of the paper is published

under the title '**Reasons**' in J Coleman and S Shapiro (eds), *The Oxford Handbook of Jurisprudence and Philosophy of Law* (Oxford: Oxford University Press 2002)

© John Gardner and Timothy Macklem

The typescript appears here with the consent of the publisher, under the publisher's eprint policy, or by author's reserved rights. Please do not quote from or cite to this eprint. Always use the definitive version for quotation and citation.

Reasons, Reasoning, Reasonableness[†]

JOHN GARDNER* AND TIMOTHY MACKLEM**

In differentiating human beings from other animals Aristotle emphasizes human excellence in the closely connected faculties of speech and reason.¹ We may think of these faculties, in their most developed form, as the distinctively human ways of relating to the world. One, the faculty of speech, provides us with a distinctive way of imposing ourselves on the world. The other, the faculty of reason, is the distinctive channel through which the world, in return, imposes itself on us.

Of all human practices the practice of the law serves most clearly to bring to a head the question of *how* the two faculties are connected. Many of the timeless problems of jurisprudence revolve around the puzzling way in which legal speech creates legal reasons and the consequent difficulty that legal reasons are often encoded in legal speech. Recent jurisprudential fashion has played up the speech side of the equation. Starting with Hart's interest in the philosophy of language in general and the theory of speech-acts in particular, through Dworkin's engagement with the theory of interpretation, through the skeptical application of

[†] This essay was defended in whole or in part, by one or both of us, in seminars at University College London, the University of Toronto, Rutgers University, and New York University. We are grateful to all who participated. Constrained by space, we have said nothing here to allay the most far-reaching queries and objections that were raised on these occasions. We hope to have the chance to tackle some of them elsewhere. Meanwhile, we owe special thanks to Grant Lamond for his extremely testing interrogations, which forced us to come clean on many issues.

* Professor of Jurisprudence, University of Oxford

** Lecturer in Law, King's College London

¹ Aristotle, *Politics* 1253^a7ff; *Rhetoric* 1355^b1ff.

literary theory to the law by Fish and others, and culminating in the current rebirth of interest in legal aesthetics and legal form, the study of legal reasoning has increasingly come to be regarded as subsidiary to the study of legal discourse. Partly this has been a reflection of a more general change in cultural mood, a growing sense that reason *per se* has somehow let us down, that rationalist optimism had held out a false promise of progress. But in part it was also an understandable reaction to the scientific aspirations of nineteenth- and early twentieth-century jurisprudence, which rightly struck many as having suppressed the creative, world-constructing dimension of legal life in favour of an arid preoccupation with total rationalization.

This trend in the philosophy of law mirrors philosophical fashion more generally. Although there has been a good deal of important recent work on reasons and rationality, the subject has found itself swimming against the stronger current of interest in language and interpretation.² Like other fashions, however, this one too often feeds off the most preposterous, stilted version of the fashion that came before it. Rationality is routinely represented, therefore, as a false prison of our own construction. Sometimes it is as if the idea itself is being burdened with all the elementary errors of all those who ever sought to understand it.³ Various forms of skepticism about rationality flourish in this climate. What the skeptics set themselves against is often a bizarre caricature unrecognizable as a human faculty, let alone as a human excellence. If rationality were as the skeptics say it is we would have every reason to doubt its importance to our self-

² The interpretivist turn can be glimpsed, for example, in the tendency to read claims that values are incommensurable as claims about the inconsistency of rival interpretations of the world, rather than as the kinds of claims about rational force set out in §7 below. For criticisms of this general interpretative turn in practical philosophy see e.g. Michael S. Moore, 'The Interpretive Turn in Modern Theory: A Turn for the Worse?', *Stanford L Rev* 41 (1989), 871; Joseph Raz, 'Morality as Interpretation', *Ethics* 101 (1991), 392.

³ For a subtle negotiation of this tendency as it figures in feminist thinking see Genevieve Lloyd, *The Man of Reason* (2nd ed., London: Routledge 1993).

understanding as human beings. But fortunately rationality is not as it is represented either by its most incautious enthusiasts or by the detractors who make capital out of that incaution, and properly understood its significance is still every bit as inescapable as Aristotle held it to be.

In this paper we will highlight some aspects of the role of reasons in human life that have often been misunderstood. Some of the misunderstandings came of the exaggerated expectations of rationality's enthusiasts, while others came of skeptical over-reactions to that enthusiasm. We begin by reasserting, against the skeptics, the classical idea that as rational beings we are beings in the world responding to the world (§1). Responding to reasons is responding to facts, not responding to one's own grasp or one's own construction of the facts. But as the essay goes on certain widely-touted supposed implications of this view that reasons inhabit the realm of fact are refined, doubted, or flatly denied. The much-misunderstood contrast between fact and value is nuanced (§2), allowing a clearer view of the relationship between reasons and values (§3). The sense in which, in spite of their facticity, reasons can somehow be personal to each of us is explained (§4). The rationalistic urge to suppose that in view of the importance of reasons everything should be done by reasoning with those reasons is deflated (§5). Steps are then taken to alleviate the anxiety that reasons are fundamentally constraining, with a broadside against the analysis of reasons as mandatory (§6) and an explanation of the importance of incommensurability (§7). Finally, in the light of all these manoeuvres, the contrast sometimes drawn between being reasonable and being rational is subjected to critique (§8). The net result is to reaffirm that the special kind of responsiveness to the world that lies in being a rational being is not all there is to being a human being, even though one is not a fully-fledged human being without it. Rationality itself also makes space for the speaking, creating, constructing agent.

By and large our examples are not drawn from the law and the specific attention we give to legal reasons and legal reasoning is extremely limited. We are among those who believe that, in spite of its important place at the interface of speech and reason, there is no special legal mode of rationality nor any special legal linguistics. Legal thought and action is subject to the same fundamental doctrines and principles of rationality (and legal utterance to the same fundamental rules of language) as the rest of human life. This view will not be defended here, but to justify the place of this essay in this book it must be assumed.

1. REASONS AND FACTS

A solitary man leaves the darkness of the Embankment and hurries onto Waterloo Bridge, looking over his shoulder repeatedly. Reaching the middle he stops suddenly, looks first behind, then in front of him and makes a lunge for the river. Caught by a passer-by at the last moment he struggles for freedom, crying out that he is being pursued from all sides by men with machetes. In fact he is not. Does he nevertheless have the reason that he says he has to jump into the river? If men with machetes were indeed pursuing him he would have that reason to jump into the river, assuming only that his chances against the river are better than his chances against the machetes. Does the fact that no-one is pursuing him mean that he doesn't have the reason to jump that he thinks he has? Is his belief still a reason even though deluded?

A year later, after successful treatment for paranoia, the man (call him A) is reunited with his rescuer. He thanks the rescuer profusely, apologizes for the trouble that he caused, and explains that at the time he was under the impression that he was being chased by men with machetes. In fact he had no reason to jump and every reason not to. He was ill at that stage and acting irrationally. Is he right to so denounce the reason he gave at the moment of his rescue? Or did his belief, that he was being

pursued to the point of death, however deluded it may have been, constitute a reason to jump into the river?

Which testimony should we trust? Concerning a particular moment in his life, the moment when he was about to jump from the bridge, A has said different things at different times. On the bridge itself he maintained that he had a reason, but a year later he said that he had none. The contradiction between these statements cannot be escaped by drawing attention to the lapse in time between the two claims. It goes without saying that A had no reason to jump into the river after his cure. The question is only whether he had such a reason at the moment when he was about to jump. At that moment he said yes, later he said no. Which statement is true?

It is tempting to decide in favour of what he originally said on the bridge. It is tempting to conclude that his belief at that time, although false, was a reason for him to do as he did. After all, only that belief *explains* what he did. And isn't it the job of reasons to explain? Yet notice that neither on the bridge nor afterwards did A himself rely on his belief as a reason. On the bridge he pointed to the presence of his pursuers, not to the presence of his belief in their presence. Afterwards he pointed to the absence of any pursuers, not (as he might if were 'in denial') to the absence of his belief in their presence. When he cited his earlier belief as an explanation of his conduct on the bridge it was just to deny that that belief was a reason for his conduct. In fact what he said was that he had acted irrationally, because his actions had been based on a deluded belief.

It might be thought that the problem here, the explanation of why the earlier belief was not a reason, is precisely that the belief in question was deluded. The poor fellow was out of his mind. Perhaps things would be different if he had been acting on a belief that had some basis in fact. Wouldn't *such* a belief be a reason? Perhaps. Take a different sort of story, then, one in which there is no delusion, no bridge, no attempt to leap into the river. Suppose that in the darkness of the Embankment, a man

sees a small knot of burly men turn onto the path in front of him, men who look very like men with whom he has just had an altercation in a pub. They are too close now for him to turn tail and retreat. If he strikes quickly, so as to take out the largest of them, he may be able to break through to a busy, well-lit road just ahead. If he does not strike he will be overwhelmed, or so he thinks. In the event he strikes, and thereby succeeds in reaching the road ahead, only to be intercepted there by a policeman.

In fact, and as the policeman soon establishes, the burly men were German tourists, hoping to enjoy what they took to be the romance of the riverside, seeking directions to the nearest Underground station from one whom they took to be a friendly local resident. Informed of this, our man (call him B) is embarrassed and extremely apologetic, but maintains that this was a genuine case of mistaken identity. The burly men may have been tourists, but in the darkness of the Embankment, not far from the pub where an altercation had taken place shortly before, their size and number gave him every reason to believe that they were they were the men with whom he had quarreled, and from whom he might well expect an attack. Does this mean that B had reason to strike them as he did?

The salient difference between A's case and B's case is that B goes on to defend his mistaken belief in terms of the reasons that he had for holding it. He was not deluded all the way down. He really did have an altercation, and the burly men who approached him really did look like the men in the pub. Although still apologetic, he does not denounce his former self-explanation as irrational in the way that A, cured of his paranoia, did. But the fact that B does not regard himself as having been irrational does not mean that he still thinks that he had a reason, in the form of his mistaken belief, to attack the German tourists.⁴ It is compatible with his making a more modest claim. It is

⁴ Compare Derek Parfit's remark that something 'not deserving the extreme charge "irrational" [may be] open to rational criticism': Parfit, *Reasons and Persons*, (Oxford: Oxford University Press 1984) 119.

compatible with his admitting that he had no such reason, but nevertheless asserting that he had reason to believe, as he did believe, that he had such a reason. His rationality lay in the fact that he responded correctly to reasons for belief rather than to reasons for action. Having responded correctly to reasons for belief he naturally, as a rational agent, responded to the reasons for action that he was led to believe he had. But still, just like A, in truth he had no such reasons.⁵

The contrast here is between having reasons for action and having reasons to believe that one has reasons for action. It corresponds to the distinction, well known to all lawyers, between justifications and excuses. One *justifies* one's actions by reference to the reasons one had for acting. One's actions are *excused* in terms of the reasons one had for believing that one had reasons for action. To be exact, the paradigm excuse is that one had a justified belief in justification. This excusatory case is sometimes known to lawyers as a case of 'putative' or 'perspectival' justification.⁶ The qualifications 'putative' and 'perspectival' show that something less than a real justification for the action is involved. This means, to our minds, that something less than a real reason for the action is involved. Instead it is a real reason for belief that there is a reason for action.

We should make clear that not all excuses are belief-based. Some are based on the argument that one acted unjustifiably on the strength of justified emotions such as anger and fear. These

⁵ In one idiom A and B alike had *explanatory* (or *motivating*) reasons but not *guiding* (or *normative*) reasons. Although this idiom has its uses we have avoided it here in favour of the idiom according to which A and B lacked reasons for their actions but regarded themselves as having them, this regard being the explanation for their actions. To put it simply, as we present it all reasons are guiding reasons. For the other way of speaking, see Michael Smith, *The Moral Problem* (Oxford: Blackwell 1994); also many of the contributions to Garrett Cullity and Berys Gaut, *Ethics and Practical Reason* (Oxford: Clarendon Press 1997).

⁶ See e.g. Suzanne Uniacke, *Permissible Killing* (Cambridge: Cambridge University Press 1994);

are adaptations of the same basic idea, however, the idea that from the inside of one's anger, or fear, as from the inside of one's erroneous belief, actions seem justified which in fact are not. Nevertheless the anger or fear or erroneous belief may itself be justified, and that is the ground of one's excuse.⁷ Such excusatory cases, of which B's is an example, differ from cases like A's. A is not responsible for his paranoid actions, with the result that those actions do not call for justification or excuse, for the expectation of justification and excuse applies only to responsible agents.⁸ This is what A is trying to convey when he apologetically denounces his deluded actions as having been performed in the grip of irrationality. By contrast B is asserting his responsibility when he apologetically pleads his mistaken but justified belief as an excuse for his admittedly unjustified action. In defending his actions he has no wish to see his rationality impugned, even though he admits his error.

These thoughts tend to corroborate the claim that reasons for action are facts rather than beliefs. Or at any rate, they lend support to the proposal that *false* beliefs are not reasons for actions, even if they are beliefs which one has every reason to hold. Neither B nor A had the reasons that at the time of their action they believed themselves to have, because their beliefs departed from the facts. Admittedly this does not quite eliminate the possibility that reasons for action are beliefs rather than facts. Perhaps reasons for action are *true* beliefs.

Consider a new example, then, one in which reasons are present but corresponding beliefs are absent. Suppose that a young woman, newly appointed to a position as a television presenter, has had the misfortune to attract the interest of a

⁷ John Gardner, 'The Gist of Excuses', *Buffalo Crim L Rev* 1 (1998) 575.

⁸ Notice that it does not follow from this proposition that A is beyond the pale of justification and excuse right across his life. He is beyond the pale only in respect of those actions performed on the strength of his paranoid delusions. On this point see Anthony Kenny, *Freewill and Responsibility* (London: Routledge 1978), 80-84.

celebrity stalker, a man who is preparing to kill her. She knows nothing of this stalker, let alone of his murderous intentions. The police however, have learned from their sources all about this man and his intentions, and they have been told that there is a letter bomb on its way to her. Having failed to intercept the bomb, do they have reason to warn the young woman, call her C, not to touch the post until they reach her house?

Of course they do. Who would deny it? But exactly what reason do they have? It is part of the very idea of a warning, like the idea of advice, that it draws another person's attention to the reasons that, according to the warner or adviser, the other person *already has*. For what they do to count as a warning, therefore, the police must be drawing C's attention to a reason that, according to them, C already has not to open her post. The reason is that there is a bomb in it. Of course C is unaware of this reason. The police know of it and she does not. But if reasons were (true) beliefs then C would have no reason not to open her post until she knew (and hence truly believed) that there was a bomb in it, and so would have no reason not to open her post until she was warned not to do so. From this it follows that her reason not to open her post could not possibly be a reason for the police to equip her with that knowledge. In other words, they have nothing to warn her of. This conclusion is bizarre. It follows that the fact that C does not yet have the belief that she is in danger cannot be an obstacle to her having a reason to avoid that danger. The reason (for C not to open the post and for the police to warn C not to do so) is the fact that there is a bomb in the post, not C's true belief in that fact, a belief which, at the time when it is needed to justify the warning, C does not yet possess.

So reasons, it appears, are facts not beliefs. One of the most common objections to this view comes of the connection that we already drew attention to, between reasons and justifications. If C has the same reason whether or not she knows of it, wouldn't it follow that all else being equal she would be justified

in not opening her post even if blissfully unaware of its contents? This would suggest that one is justified whenever one accidentally engages in the course of action that reason would support, irrespective of one's reasoning. This conclusion seems too much to bear. Fortunately we have no need to bear it. There are at least two independent conditions that must be satisfied to establish that one was justified in what one did. First, one must show that reasons did indeed support the course of action that one has engaged in. Secondly, one must show that one acted *for* one of those supporting reasons.⁹ It follows that C is not justified in declining to open her post if she is not aware of any reasons not to, for in that case she cannot act *for* such reasons. Nevertheless the reasons exist and are capable of contributing to the meeting of the first condition, whether or not she is aware of them. So the fact that unknown justifications are not justifications is in no way incompatible with the claim that reasons are facts, not beliefs.

Those who are committed to the opposite idea, that reasons are beliefs, and yet who find that conclusion unpalatable with respect to certain beliefs, particularly with respect to the paranoid belief held by A (that he was being pursued by men with machetes), must exclude certain beliefs from the realm of reason without appealing to a relationship between those beliefs and the facts. In order to achieve this, and so save their position, they need to appeal to the quality of those beliefs *as beliefs*. They need to appeal to what we might call the 'beliefness' of those beliefs, where this quality does not lie in the relationship between those beliefs and the facts, that is in their facticity. Can they do so? The very contrast between beliefness and facticity is an odd one, since it is in the nature of a belief that it aims at the facts, so that it might be thought that whatever beliefness is, it is necessarily proportionate to facticity. The thought, we suggest, is well

⁹ The second requirement is defended in John Gardner, 'Justifications and Reasons', in A.P. Simester & A.T.H. Smith (eds.), *Harm and Culpability* (Oxford: Clarendon Press 1996).

founded. In truth what these adversaries of ours need to appeal to in order to identify a reason (or absence of a reason) are facts (or the absence of facts), and accordingly the arguments that they actually offer are plausible just to the extent that those arguments make an admitted or tacit appeal to the relationship between beliefs and facts, that is, to the facticity of the beliefs.

In spite of elaborate attempts to show that there is something wrong with A's beliefs independent of the facts, such that A does not have the reason to jump that he takes himself to have, in the end the only way to isolate an error in A's reasoning is by reference to the facts. Is there anything else wrong with A's beliefs? Are they incoherent with his other beliefs, or does he fail to believe that his beliefs are justified, to pick two criteria that have sometimes been suggested as tests for beliefness? Not a bit of it. His disorder is not a form of dementia or compulsion. He is not confused within himself, or reluctantly struggling against alien impulses. On the contrary, he is at ease with his consistent but deluded view of the world. His belief in the presence of men with machetes is irrational, not because it is incoherent or alien, and hence deficient in beliefness, but just to the extent that it is deluded, that is, remote from the facts. The allure of appeals to coherence, or to second-order cognitive endorsement, or to any other relations among beliefs, comes of the tacit links between these tests and authentic tests of facticity. People are lulled into thinking that coherent or endorsed beliefs are more likely to be true (and *vice versa*) and it is on the basis of that truth that they are drawn to such explanations of rationality. But they should not be so lulled. A's lack of rationality relative to B lies in a lack of connectedness to the facts. The real reason why B is so anxious to distance himself from the accusation of irrationality is that he does not admit to being that detached from the facts and therefore from the world. The gist of B's excuse is that he was very nearly fully connected to the facts. For reasons are facts.

2. FACTS AND VALUES

It may be said, even by those who agree that reasons are facts as opposed to beliefs, that this only tells half the story. A *complete* reason is actually made up of two components, it may be said, and only one of these is the admittedly factual component. The other component is *evaluative*. In answer to a request for reasons we sometimes state only the factual component (that there is a bomb in C's post) but at other times state only the evaluative component (that C's being blown up by a bomb would be terrible). Which of these we choose to state and so emphasize depends upon which of them our interlocutor can be expected to already regard as obvious, so as to go without saying. Sometimes both components are complex and call for expression. But typically one will suffice. Nevertheless, it is said, the other component always lurks in the background.¹⁰

We agree that many statements of reasons are incomplete because they take certain things for granted. Suppose that D is asked by a friend whom she meets in the wine shop why she is buying a bottle of Trebbiano. Suppose that D's answer is simply that she is having spaghetti vongole for dinner. Her friend is baffled. Is Trebbiano an ingredient in spaghetti vongole, he asks? No, says D, but it is a spectacular accompaniment to that dish. This response makes clear that her original answer was not a complete statement of her reason for buying Trebbiano, or to put it another way, did not state a complete reason for buying Trebbiano. Her complete reason has two components, one

¹⁰ The most common version of the view according to which every complete reason cannot but have two components is the view according to which every reason comprises a pro-attitude and a belief. This view is owed first and foremost to Donald Davidson, 'Actions, Reasons and Causes', in Davidson, *Essays on Actions and Events* (Oxford: Clarendon Press 1980), and has been adopted by many others, including, for example, Philip Pettit and Michael Smith, 'Backgrounding Desire', *Philosophical Review*, 99 (1990) 565.

operative and the other *auxiliary*.¹¹ The operative component of her reason is that Trebbiano is a spectacular accompaniment to spaghetti vongole. The auxiliary component is that she is having spaghetti vongole for dinner. Together these components add up to a complete reason for her to buy Trebbiano. It is not the only possible (complete) reason to buy Trebbiano, nor the only possible reason to buy Trebbiano that contains the auxiliary component that D is having spaghetti vongole for dinner. As her bemused friend's question revealed, it might be a reason to buy Trebbiano that according to some recipes Trebbiano is an ingredient of spaghetti vongole rather than an accompaniment to it, and that D's partner has decided to make spaghetti vongole according to one of those recipes. Similarly the same operative premiss may be combined with a variety of different auxiliary ones. The fact that Trebbiano is a spectacular accompaniment to spaghetti vongole might also be a reason to *make* spaghetti vongole, if D and her partner have not yet decided on the menu for dinner and Trebbiano is what they are drinking tonight.

But is it true that every complete reason comprises both these components? Consider the operative premiss that D mentions by way of further explanation for her choice of wine, namely, that Trebbiano is a spectacular accompaniment to spaghetti vongole. Isn't this a complete reason for a different action from the one that D is performing, namely, to try the combination of the two? It would not be a complete reason for buying Trebbiano or for having spaghetti vongole, for the reason for those actions is incomplete in the absence of an auxiliary component, that D and her partner are having spaghetti vongole tonight, or have already opened a bottle of Trebbiano. Yet the operative premiss is indeed a complete reason for trying the combination of the two and can be recognized as such by D's friend. For example, that premiss is all that D's friend needs to offer by way of explanation

¹¹ We borrow this terminology from Joseph Raz, *Practical Reason and Norms* (2nd ed., Princeton, NJ: Princeton University Press 1990), 33–35.

the following week when he is asked by his partner why he is so insistent on this particular combination of food and drink. Doesn't this show that while every auxiliary premiss needs an operative premiss to complete it as a reason the reverse does not hold true?

A natural response to this contention might be to try to show that even this reason (that Trebbiano is a spectacular accompaniment to spaghetti vongole) breaks down into two components, that what we have paraded as an operative premiss is in fact an operative premiss and an auxiliary premiss combined. To state that Trebbiano is a spectacular accompaniment to spaghetti vongole, it might be contended, is to incorporate by reference an unstated operative premiss, that spectacular combinations of food and drink are always worth trying, relative to which the stated and supposedly complete reason is strictly speaking only an auxiliary premiss. But this response fails. The new suggested operative premiss is vacuous. Not only does it go without saying that spectacular combinations are always worth trying, but the reason that it goes without saying is that the idea that they are always worth trying is built into the idea that they are spectacular. In other words the proposition is analytic. Having said that Trebbiano is a spectacular accompaniment to spaghetti vongole, one adds nothing at all by way of further explanation of why one is trying that combination, let alone anything *operative*, by adding that spectacular combinations are worth trying.

Perhaps it is possible to find some other ingenious way of breaking down the complete reason that Trebbiano is a spectacular accompaniment to spaghetti vongole into two components, one genuinely operative, the other genuinely auxiliary. All that we are saying is that a point will always come at which further dismantling of successive operative premisses introduces a vacuity. At that point the vacuous pseudo-operative premiss drops out, what was presented as the auxiliary premiss

turns out still to be operative, and that operative premiss is a complete reason in its own right.

This is not to diminish the rational importance of auxiliary premisses. It leaves their central role intact, which is their role in practical reasoning. Their role is to transmit the force of one complete reason for action to another action, namely an action that will contribute to the performance of the first. For example, buying Trebbiano contributes to trying the spectacular combination of Trebbiano and spaghetti vongole just in case one is already planning spaghetti vongole for dinner. Likewise A's jumping into the river contributes to his escaping men with machetes just in case there are men with machetes in pursuit of him and the river will carry him away from them at less peril to himself. In these cases, as in all others, the auxiliary premiss is a reason for action given the operative premiss. It is a reason for performing the action which in fact will contribute to the performance of the further action for which the operative premiss is a complete reason. That is why explanations citing auxiliary premisses, such as the explanations offered by A, B and C, are capable, if the auxiliary premisses are true, of being rational explanations.

At this point it might be said that we have cast doubt on our own claims in the previous section. We maintained there that reasons are facts. We then started this section by raising a challenge to that view, namely, the challenge that facts tell only half the story. What they tell is the auxiliary half of the story. The operative half of the story, the challenge continues, is evaluative rather than factual. In arguing that operative premisses can be complete reasons without relying on auxiliary premisses haven't we confirmed that some reasons do not have factual components? In the case of such reasons it is surely not even a *half* truth to say, as we said, that reasons are facts?

In endorsing the distinction between operative and auxiliary premisses, however, we did not endorse the view that operative premisses are not facts. What is true is that operative premisses

are evaluative, but they are evaluative precisely because the facts that go to make them up are value-laden facts. They are facts like the fact that Trebbiano is a spectacular accompaniment to spaghetti vongole, or the fact that the Lake District is beautiful, or the fact that the Thames is dangerous. The elemental error of those who try to break down every complete reason into operative and auxiliary components is the error of thinking that while there may be mixtures of facts and values there are no true compounds of the two. It is always possible, it is said, to separate the factual from the evaluative without changing the character of either and without slipping into vacuity. Yet one need not believe in the strong thesis that all evaluative properties are supervenient upon factual ones to deny that such separation is always possible.¹² Either way, holding that all reasons are value-laden is in no way incompatible with holding that all reasons are facts, through and through.

3. REASONS AND VALUES

So it seems that all reasons are value-laden. Every complete reason contains an operative component which reveals the value that would be served by the action for which it is a reason. If this is true it raises the question of the relationship between reasons and values. In particular it may bring to mind a contentious thesis that we could call *proportionalism*. Proportionalism is a thesis that holds the relationship between reasons and values to be constant. According to proportionalism, to be exact, one always has more reason to perform the more valuable of any two actions. Some proportionalists hold that this doctrine is analytically true,

¹² On some varieties of supervenience relevant to practical reasoning see R.M. Hare, 'Supervenience', in Hare, *Essays in Ethical Theory* (Oxford: Clarendon Press 1989), or John McDowell, 'Non-Cognitivism and Rule-Following', in McDowell, *Mind, Value and Reality* (Cambridge, Mass: Harvard University Press 1998).

because statements of reasons are simply synonymous with statements of values. In other words they define reasons as values or values as reasons.¹³ Other proportionalists hold that this doctrine is a synthetic truth, in that reasons track values or values track reasons even though the two are not the same thing. Sometimes proportionalist views have been associated with consequentialist moral theories, or more broadly, with instrumentalist views of rationality.¹⁴ As our formulation of it reveals, however, proportionalism need not be such a parochial creed. A proportionalist may believe that reasons track values or that values track reasons. The latter possibility is the one associated with deontological moral views, views according to which, in the Rawlsian idiom, the right is prior to the good.¹⁵ So long as the values in question are held to track the reasons consistently, or in other words so long as actions are held to be good just to the extent that they are right, such deontological moral views are fully compatible with proportionalism.

Nevertheless many people do take issue with proportionalism. Their common anxiety tends to be this one.

¹³ We say 'or' because the direction of definition typically depends on which of the two ideas the readership is deemed already to understand. In *What We Owe to Each Other* (Cambridge, Mass: Harvard University Press 1998) T.M. Scanlon 'passes the buck' from the concept of value to that of reason. Others may regard it as more perspicuous to 'pass the buck' in the other direction. We doubt whether this buck should be passed at all, not because there can be no valid interdefinition of values and reasons, but because neither of the two concepts really is any easier to understand than the other.

¹⁴ See John Finnis, *Fundamentals Of Ethics* (Oxford: Clarendon Press 1983), 85-86. We think that what we call proportionalism is the same thing that Finnis calls by that name.

¹⁵ Rawls in fact divides deontological moral views into two types. A deontological theory 'either does not specify the good independently from the right, or does not interpret the right as maximizing the good': *A Theory of Justice* (Revised ed., Oxford: Oxford University Press 1999), 26. Possibly deontological theories of Rawls' second type do defy proportionalist interpretation, although, for reasons that will emerge in §5 and §6 below, the matter is not straightforward.

Surely, they say, values are impersonal things, whereas reasons are (or can be) personal.¹⁶ I can have my reasons and you can have your quite different reasons, without any implication of rivalry. But my values and your quite different values, the thinking goes, can only be rivals. If reasons are personal and values are impersonal, it is concluded, then reasons cannot possibly correspond to values. They must vary somewhat independently of each other. The variations could be in either direction. There could be reasons that do not correspond to values. There could also be values that do not correspond to reasons. Different critics of proportionalism emphasize one or other of these possible asymmetries. But are any of these critics right? Are there any genuine counter-examples to proportionalism?

a) Reasons Without Values

One attack on proportionalism cites the possibility of there being reasons that do not correspond to values. The operative components of those reasons, it is said, do not, strictly speaking, imply corresponding positive evaluations of the actions for which they are reasons. The most common suggestion, and the one that we will emphasize, is this. Desires as well as values can constitute reasons for action.¹⁷ Perhaps so, the proportionalist could

¹⁶ For a survey and exemplification of this anxiety see David McNaughton and Piers Rawling, 'On Defending Deontology', *Ratio* 11 (1998) 37.

¹⁷ A different kind of counter-example to the view that all reasons correspond to values might be thought to arise in the law. Legal reasons are source-based rather than merit-based. In other words, if asked for a legal reason for doing something the correct response is always to cite some authority rather than some value. But this is no counter-example. Whenever a legal authority creates a legal reason to do something, doing that thing is claimed by the authority to be of value. Does this mean that sometimes reasons correspond only to *claimed* rather than actual values? It does not, for to the extent that the values are only claimed rather than actual the legal reasons are only claimed to be reasons, that is, they are only reasons from the legal point of view. For one explanation of this phenomenon see John Gardner, 'Law as a Leap of Faith',

respond, but even if desires figure in reasoning they need not be severable from values in order to do so. First, the fact that something is desired could itself be one of the things that lends value to its pursuit. Secondly, desires themselves could answer to values, in which case the appearance that a desire is a reason could be explained away by pointing to the values that support the desire. Either way the introduction of desires into our picture of reasons would drive no wedge between reasons and values.

Are either of these suggestions about the role of desires in rationality tenable? Could there be value in the fact that desires are satisfied? Or could desires answer to values? We think that both suggestions are tenable, but that they are interlocked, so that the first possibility depends on the second. When desires are sufficiently supported by reasons they are themselves capable of becoming further reasons. When that condition is met, they are capable of becoming our *goals*, and goals do not merely reflect reasons but also constitute them.

Think about a case in which the condition is not met. Suppose that a man, call him E, checks to make sure that he has locked his front door before leaving for the office. Having checked the door he starts down his front path, only to turn at the gate and check the door again, this time pushing against it heavily several times. Having checked the door a second time he walks down his front path and gains the street, only to turn and check the door yet again. And so on, over and over, so that it takes him ten minutes to leave the house finally, still looking over his shoulder as he turns the corner of the road. Is his desire to check the door on each successive occasion a reason to check the door on that occasion? We can discount the reason to check the door so as to avoid the anxiety that he may otherwise feel during the course of the day as the result of his desire remaining unsatisfied. This reason corresponds to a value, the value of

having a day that is free from anxiety, for a day of anxiety is a day when other valuable things could have been achieved, as well as a day of negative experiences. E might cite his desire as the reason here but if he did he would be suppressing the operative premiss, namely that nagging desires drain value from one's day. The fact that E has the desire is the auxiliary premiss of this reason. Our question is whether the desire to check the door on each successive occasion is any reason to check the door once the removal of anxiety has been subtracted from the story. Is the desire is ever operative, and so a reason in its own right?

The case is a variant on that of A, whose action was based on a pathological belief and who did not have the reason that he thought he had to jump from the bridge. E's action is based on a pathological desire, which is in turn based on a pathological belief that he has left the door unlocked. Is there any more reason to check the door than to jump from the bridge? E is unlikely to say that he just felt like checking the door, without giving a further reason. Even if he would say such a thing it would make him doubly crazy. Who just feels like checking a door, that they know full well to be locked, in order to make sure that it is locked? What E is more likely to say is that he was unsure whether the door was locked, that he could not remember whether he had locked it, that he had been robbed once as the result of leaving his door open and wanted to be sure not to leave the door open again, or something like that. All these would be good reasons to check the door, and to want to check the door, were E not so palpably deluded in his belief that he had left the door unlocked.

E is likely to offer such an explanation just because he recognizes that desire, like action, answers to reason. Where desire is unsupported by reason, as the inadequacy of E's explanations shows that it was in his case, desire is itself no reason to do what would satisfy it. As E recognizes, his desire would only be a reason for action if it were either a valuable desire (i.e., if it really did contribute to reducing the risk of his door being

unlocked), or if it were rendered rationally significant by the value of avoiding the frustration or anxiety that would come of its non-satisfaction (in which case it would be auxiliary to that operative premiss). Either way it would fail to disrupt the proportionate relationship between reason and value.

Perhaps, as in A's case, the problem here is that the desire is a pathological one. Perhaps more compelling counter-examples could be gleaned from consideration of the apparently paradoxical cases in which people deliberately set themselves, or claim to set themselves, against the pursuit of value. These cases divide into three types. First there is the simple case of people who set themselves against the pursuit of value without realizing what they are doing, for they mistakenly hold to be valuable that which is not, and not valuable that which is. Such people call for no reassessment of the relations between reason and value. Their situation is akin to that of B. In so far as they are mistaken about value, they are likewise mistaken about the reasons for action that apply to them. In so far as they are nevertheless rational this is because they nevertheless have reason to believe what they mistakenly believe. At best they are excused rather than justified in what they do on the strength of such a belief.

The second type of case is more problematic. It is that of *akrasia*, or weakness of will. This is often thought to pose the greatest challenge to rational accounts of human agency, for it is a case in which desires for one action overcome the reasons that the agent is aware would lend more support to the performance of another.¹⁸ But this characterization itself makes clear why *akrasia* poses no challenge to proportionalism, for in so far as the agent knowingly acts against reason she knowingly acts against value. Or at any rate there is nothing in *akrasia* to suggest otherwise. In effect it is another, albeit less dramatic and more everyday, case of the pathological. It is less dramatic because in

¹⁸ The *locus classicus* is Donald Davidson's 'How is Weakness of the Will Possible?' in his *Essays on Actions and Events*, above note Error! Bookmark not defined..

general *akratic* people have a reason to do what they do, and correspondingly there is value in it. The problem is only that in the light of the alternatives they do not have *enough* reason, *enough* value, to justify what they do, as well they know.

This leaves us with the third type of case. This is the type of case in which the agent *literally* sets himself against the pursuit of value. What he aims at is evil *qua* evil.¹⁹ There is of course no need for us to show that this character has reasons for what he does. Perhaps he is irrational. But how is it even intelligible that he thinks as he does? How could he imagine, if proportionalism is true, that he could have reason to pursue evil without pursuing it under the description of countervailing good? Must we conclude that proportionalism is false, that there are things other than values that could conceivably count as reasons? Not so. In so far as this character is intelligible he does indeed secretly pursue (what he regards as) evil under the heading of good. As the ridiculous pursuits of Satanists the world over demonstrate the description evil is a kind of code-word for what is covertly regarded by its pursuer as a good thing. Perhaps it is the value of membership in a group sharing the *frisson* of forbiddenness, or perhaps it is just the value of contrariness. One should not deny that there may be real value in these things. Of course this real value, in so far as it is a value at all, may not be enough to justify the pursuits in question when set against their genuinely bad features. But this only goes to show that proportionalists have nothing to be afraid of in this kind of example. The question of what would justify these pursuits is indeed a question of their value, for one has no other reasons to cite in defence of them than those which correspond to actual values that they have.

¹⁹ Sometimes deliberate infliction of pain or suffering is thought to fit this anomic profile, but in most cases it does not. Normally it is thought to be deserved, or justified for the greater good, and valuable under that heading. The view we offer on these matters is akin to that of Warren Quinn in 'Putting Rationality in its Place', in Quinn, *Morality and Action* (Cambridge: Cambridge University Press 1993).

It is tempting to conclude from this that desires drop right out of the rational picture, that they do not add to the case for doing anything, but provide people at best with motivational fortification (and at worst with motivational obstacles) in doing what there is in any case reason for them to do. But this would be an overly hasty conclusion. When one's desires are sufficiently supported by reasons the fact that one has these desires can give one further reasons to do as one desires. These further reasons represent the value of one's being personally engaged, via one's desires, in what are in any case valuable pursuits. Can this explain the idea, flagged earlier, that reasons for action are somehow personal things, that I have my reasons and you have yours?

It does not yet provide a full explanation. One's engagement is valuable, and this value, like other value, is in principle there for all to pursue. There is no suggestion in what we have said so far that the value of personal engagement is an agent-relative value, one that figures especially in the life of the engager and is thus especially relevant to his or her practical reasoning.²⁰ In so far as it yields reasons it yields reasons not only for each to be personally engaged but also for all to contribute to that personal engagement. Thus if there is a sense in which some practical reasons are personal ones, this remains to be explained.

b) Values Without Reasons

We conclude that if there is a case against proportionalism it does not lie in the existence of reasons that do not correspond to

²⁰ The distinction between the agent-relative and the agent-neutral was introduced by Thomas Nagel in *The Possibility of Altruism* (Oxford: Clarendon Press 1970) 90-95 and named in these terms in his 'The Limits of Objectivity', in Sterling McMurrin (ed.), *The Tanner Lectures on Human Values, Volume 1* (Salt Lake City: University of Utah Press 1980) 79, at 102. The term 'agent-relative' has often been used in looser ways, for example to denote special responsibilities. However in the strict sense special responsibilities are agent-neutral if there are reasons not only for A to fulfil them but also for B, C, D etc. to contribute to A's fulfilment of them.

values, for there are none. Does it lie in the converse possibility? Are there values that do not correspond to reasons? Take the case of two friends who haven't been getting along with one another recently. Recognizing the threat to their friendship they agree to go out for the day to see if they can mend matters. One of them, overly anxious about the situation, does everything he can to please the other, much to the other's annoyance. A brief argument between them ensues and the anxious friend, call him F, promises not to try to please his companion further. Of course, if he keeps this promise part of the value of his doing so will be that it pleases his friend. Yet if F acts *in order to* please his friend, if that value is operative in his reasoning, he will break his promise. It follows that the value of pleasing his friend gives F no reason to keep his promise, despite the fact that this value is part of the value of keeping the promise. F has every other reason to keep his promise, but not this one.

The explanation of this phenomenon lies in the fact that it is part of the nature of a reason that it must be logically possible to do as that reason would have one do *by acting for that reason*. The problem in F's case is that if he tries to keep his promise in order to please his friend he necessarily breaks his promise. It would be logically self-defeating for him to act for that reason. It follows that he does not have that reason to act, even though his action has the corresponding value, that is, the value that would correspond to the reason if he had it.

It does not follow that this value gives *nobody* any reason to act. Suppose that F has another friend G, who is aware of the promise and of how much it matters to F's friendship with the promisee. G also knows that F is a very anxious person, and that it will not take much, perhaps only a comment on the promisee's dark mood, to restore F's self-defeating eagerness to please. In order to save F from breaking his promise in this way G, who has joined the two friends for lunch, keeps silent when he might otherwise have spoken. G's reason for his silence is to forestall the promisee's displeasure at the breaking of the promise. He has

no wish to make the promisee's mood any darker. There is nothing logically self-defeating about G's acting for this reason. He can contribute to F's keeping his promise for a reason which is not a reason for F to keep that promise himself, even though it corresponds to part of the value of F's keeping the promise. To treat the fact that the promise-breaking will displease the promisee as a reason for F's promise-keeping is logically self-defeating for F, but not for G.

We doubt whether any value-bearing facts are free of rational significance for everyone. Their lack of rational significance is always relative to particular people and particular actions by those people. The most common situation is that the first party, the one whose action will carry the value in question, is the one for whom that value is no reason, whereas third parties, those who can contribute to the performance of the value-bearing action by the first party (for example, by scheming, cajoling, compelling), have rational access to that value. They can scheme, cajole or compel for that reason even though the first party cannot perform the action for that reason.

We called this the most common situation but some will say that the situation we are describing is not common at all. Some will say that our example was an unusual and contrived one, not far from the paradoxical case of a promise not to keep one's promises. We think, on the contrary, that our case is far removed from this paradoxical case. Ours is a promise of an intelligible, valid, and we think fairly common type. More important, our promise brings into sharp relief a feature of practical life that may also arise even without the intervention of any promise, command, request or similar content-independent device for creating reasons.

Consider a friendless soul, call her H, who hears everyone talking about the great value of having friends. Never having participated in this unfamiliar value she sets out to do so. She is an extremely good mimic. She quickly gets the hang of acting towards people in seemingly friendly ways and quickly picks up

what she takes to be a circle of friends. Not only that, but these people also regard her as part of their circle. All the time, however, her eye is only on the value of having friends. She never thinks about her so-called friends in the way that friends do. Their joys and sorrows are never her joys and sorrows, their passions are merely embarrassments to her, and so on. She thinks about them as beings who satisfy her need to have friends. What she does not realize is that this way of thinking about them, this way of reacting to them, is itself incompatible with her having friends. She has no friends while she aims, in her relations with her pseudo-friends, at having them as friends. Imagine H being asked, by an outsider to the circle, why she is friends with members of the circle. She says that it is in order to have friends. This only shows that she is not friends with them at all. Her pursuit of the value of having friends is logically self-defeating. Of course, H's basic problem is that she misunderstands the value of having friends. Some will say that this example is one of evaluative error rather than self-defeat in the pursuit of genuine value. But our point is that H's evaluative error lies precisely in her thinking that one can have friends in order to have friends. It is true that having friends is part of the value of friendship. But that is not a reason for performing any of the constituent acts of friendship.²¹

In our view, the example can be replicated. There are other values that one cannot in logic bring to one's actions by pursuing them. We think, for example, that the intrinsic values of some or all of the moral virtues are among these values. We suspect that there is no courage in actions performed in order to display courage, no compassion in actions performed for the sake of being compassionate, etc. These are tricky cases and we cannot develop them here. But the cases that we did develop already

²¹ Notice again that this inhibition does not carry over to third parties. It is open to G to take action for the sake of F's having friends, albeit misunderstanding that value exactly as H does. This is not logical anathema to G's helping to forge fine new friendships for F.

show that proportionalism is an oversimplified doctrine. Many writers have tried to cast doubt on the doctrine by showing that some reasons do not correspond to values. We rejected this critique but endorsed the less widely supported critique that, relative to some agents and some actions, some values do not correspond to reasons. Some values that would lie in those agents performing those actions do not correspond to reasons for those agents to perform those actions, even though they may well yield reasons for others to contribute to such performance.

4. REASONS AND PERSONS

We have started to see, through studying the relationship between reasons and values, how reasons could be more personal than values. Given one value which both F and G could serve through their actions, namely the value of pleasing the promisee, G (the third party) had a reason corresponding to that value but F (the promisor) did not. But so far this seems to be an asymmetry in the wrong direction. One would have expected the value of personal engagement to be such that in respect of his own promise F had reasons that G did not. The only glimpse of agent-relativity in reasons that we have seen so far seems, perversely, to shift the rational burden of each person's activities onto others.

But return to the story of F's promise and G's contribution to his keeping it. Consider a possible complication. Having noticed the promisee's bad mood over lunch G might be tempted to warn F to keep the promisee sweet by, for example, not breaking his promise. F's pleasing the promisee is a reason for G to contribute to that pleasing, after all, and what contribution would come more naturally than a word of friendly advice about the promisee's mood? But in the case of C and her parcel bomb we drew attention to a special feature of warnings and advice. To warn or advise someone is to draw their attention to reasons that already apply to them. We already know that pleasing the promisee is no reason for F to keep his promise. It follows that G,

knowing what F promised, cannot warn F to keep his promise for this reason. That would not be a contribution to F's keeping his promise at all. Rather it would be a contribution to his violation of it. This is another case of self-defeating action. What makes it self-defeating is not the reason for which G performs his action, but rather the action that he performs for that reason. He has the reason to act, unlike F, but the reason is not a reason to perform this tempting action, for the reason is a reason to contribute and this would not be a contribution.

This is just a special case of a very widespread phenomenon. We already drew the distinction between doing as a reason would have one do and doing it for that reason. In the last section we claimed that nothing is a reason for anyone to do anything that that reason would have them do unless it is logically possible for them to do it for that reason. Now we add a further point. Sometimes people perform actions for reasons which do not in fact support their performing those actions but which support their performing other actions instead. They have those reasons, for it would be logically possible for them to do, for those reasons, as those reasons would have them do. Unfortunately, what they actually do for those reasons is something else. For instance, they have a reason to contribute to someone else's actions but what they do for that reason is far from a contribution. It is actually a hindrance. This is an important kind of rational mistake. It differs subtly from the mistakes made by A and B, which were mistakes about which reasons apply to them *tout court*. This, by contrast, is a mistake about which actions are supported by reasons that admittedly apply.²²

²² The difference is subtle because arguably in each case the mistake lies in the auxiliary premiss. However there is a sense in which reasons apply to G and do not apply to A and B. G has reason to be frustrated if he discovers that there is nothing he can do to help F, whereas A and B have no reason to be frustrated (although they think they have) if they find they have no means of escape.

Among the most common mistakes of this kind are mistakes by third parties that ignore the value of personal engagement. This value, as we said, is an agent-neutral one like all others. Everybody's personal engagement with valuable activities is in principle everybody's business. This way of putting the point raises the spectre of a world of nosy parkers, intruding into each other's lives at every turn. It reminds one of the familiar objection to utilitarianism that it gives everyone one and the same goal of serving everyone's goals.²³ But notice that this image ignores the fact that many contributions that people might be tempted to make to other people's pursuit of their goals will not be contributions at all, for they will detract from rather than adding to the personal engagement that the very idea of pursuing a goal implies. Rather than their being more personal engagement in the world, thanks to these nosy parkers there will be less. The reason they have to contribute to other people's personal engagements are not, therefore, reasons to perform these pseudo-contributory actions.

This is the main thing that gives people the impression that reasons are somehow more personal than values.²⁴ The reality is that the reasons in question are as agent-neutral as the values (barring the special cases we dealt with in the last part of the previous section). They are as much reasons for third parties as for first parties. The point is, however, that third parties only have it in their power to make limited contributions to first-party conformity with those reasons. Thus the reasons as applied to the third parties are reasons for a more restricted and less exciting

²³ The objection is Bernard Williams' famous integrity objection in 'A Critique of Utilitarianism', in J.J.C. Smart and Bernard Williams, *Utilitarianism: For and Against* (Cambridge: Cambridge University Press 1973), 116-17. The point is further developed by Peter Railton in 'Alienation, Consequentialism, and the Demands of Morality', in *Phil & Pub Aff* 13 (1984) 134.

²⁴ We here build on Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press 1986), ch. 12. Our conclusions go beyond those that Raz defends.

range of interventions than they may be tempted to suppose. In the face of this there may be some value in people thinking agent-relatively about other people's roles and pursuits.²⁵ Of course this too can get out of hand, and can descend into an absurd egoism. Nevertheless, within limits, adopting an agent-relative perspective in our practical reasoning may help us, from day to day, in resisting the temptation to get so involved in other people's personal engagements that they stop being other people's personal engagements, and their value as such is destroyed by our well-intentioned but self-defeating meddling.

5. REASONS AND REASONING

The previous section introduced examples that illustrated the value of indirect reasoning. In these examples one had a reason to do something but in acting for that reason one was apt to do the wrong thing. Having a reason to help (say that one's friend is having trouble achieving his goals), and acting for that reason, one was apt to interfere to the point of hindering rather than helping. In such cases it is *contingently* self-defeating to act for a certain reason.²⁶ It is not logically self-defeating, because it is logically possible that when one intends to help one helps. Nevertheless features of oneself or one's predicament may impede one's success. There may be distractions, or temptations, or other complications which make it likely that in aiming to help one only hinders. This is the vice of the much maligned

²⁵ For a strong agent-neutral case for limited agent-relativism see Derek Parfit, *Reasons and Persons*, above note 4, part one.

²⁶ Parfit's *Reasons and Persons* (previous note) begins with a study of self-defeatingness. His contrast between what is directly self-defeating and what is indirectly self-defeating is close but not identical to our contrast between what is contingently and what is logically self-defeating. Parfit also emphasizes what we ignore, namely, that self-defeatingness is exacerbated by many-person activities. This is the source of familiar rational difficulties in co-ordination problems and prisoners' dilemmas.

'do-gooder', who realizes that she has reason to help others but doesn't know when to stop. She doesn't notice that at a certain point her eagerness that others should reach their goals, and consequent insistence on participating in their doing so, is precisely what gets in the way of their reaching their goals. For at a certain point *her* participation dilutes *their* participation, the very personal engagement that is part of the value of their having and pursuing goals. At this point she has a reason not to try to help. It does not follow that she has no reason to help. This would follow only if acting for the sake of helping were logically self-defeating, which in this case it is not. Nor does it follow that she has a reason not to help. What she has is a reason not to act for a certain reason, namely, for the reason of helping. Such reasons not to act for a reason have become known, following Raz, as *exclusionary reasons*.²⁷

Sometimes, obviously, it makes sense to act for the sake of protecting one's children. Getting one's children to wear their seat-belts does help to protect them from some injuries, and in the case of most people there is nothing that would be distracting or complicating in getting children to wear their seat-belts for that very reason. But other operations regarding the safety of one's children are more complex.

Consider a parent, call him J, who forbids his children to go out and play in the street in order to protect them from heavy traffic and malevolent strangers. In the short term he succeeds. But soon things start to go wrong. The older child grows up obedient but timid. When he at last finds himself alone in the streets as a young adult he is ill equipped to cope with traffic and strangers. He is unsure how to gauge the speed of approaching traffic and is no judge of character. The younger child, on the other hand, grows up rebellious and reckless. She crosses the road without regard to traffic and is utterly indiscriminating in her choice of company. As they enter adulthood both children are

²⁷ Joseph Raz, *Practical Reason and Norms*, above note 11, 39.

vulnerable to dangers that they would not have been as vulnerable to had their father J not set so much store by their safety and so forbidden them to play in the street in their teenage years. J's setting so much store by their safety, and his consequent protective action for the sake of that safety, is contingently self-defeating (assuming that it is his children's long-term safety, safety into adulthood, that he is setting store by). To do as that reason would have him do, that is, to make his children safer, it would be better for J not to have so often acted for that reason, in other words, not to have been so protective. As with so many parents, the reason tempted him into the vice of over-protectiveness. In view of that temptation he had, on at least some occasions during their upbringing, a reason not to act for that reason: an exclusionary reason.

Some people have doubted the very possibility that there can be exclusionary reasons.²⁸ One explanation for their doubt is that they misinterpret the proposition about reasons that we introduced in §3, namely, that nothing is a reason unless it is logically possible to do as that reason would have one do by acting for that reason. According to the misinterpretation nothing is a reason unless it is possible *in the circumstances* to do as that reason would have one do by acting for that reason. This misinterpretation gives rise to two myths, which we will call the rationalist myths. The combination of these two myths yields the conclusion that there are fewer reasons for action that might have been supposed but that those reasons offer greater security against error for their faithful followers.

The first is the myth often captured in the slogan 'ought implies can'. As this slogan is used by those in the thrall of the rationalist myths, it means that people who have attenuated capacities or attenuated opportunities also have, by that token, fewer reasons for action. They lack the reasons for action to do

²⁸ See for example Michael S. Moore, 'Authority, Law and Razian Reasons', *Southern Cal L Rev* 62 (1989), 829 at 859ff. This example is interesting because of its explicit subscription to the rationalist myths described below.

whatever they lack the capacity or the opportunity to do. We will not directly tackle that myth here.²⁹ Our main interest lies in the second rationalist myth. According to this myth it is in the nature of a reason that there can be no better way of doing what any reason would have one do than by acting for that reason. Reasons, the thinking goes, are for following. How is it compatible with this truism, ask subscribers to the second rationalist myth, that sometimes it is better, even so far as the reason itself is concerned, not to follow that reason? How can there ever be a reason not to follow a reason? How are exclusionary reasons possible?

One prominent development of the second myth in modern moral philosophy sees it invoked as a principle for coping with a variety of conflicting reasons. According to the second myth, recall, there is no better way of doing what any reason would have one do than by acting for that reason. For some it is a short step from this proposition to the proposition that there is no better way to act as the *balance* of reasons would have one act than by *balancing* those reasons, that is, by *deliberating* about the pros and cons of the various actions that are open to one in the circumstances.

Yet one need not introduce the idea of an exclusionary reason to show that this short step is a step in the wrong

²⁹ What we will say, however, is that one particular class of reasons is capped by capacity. These are interest-based reasons (or reasons of well-being). Something is in your interests (serves your well-being) if firstly, it is valuable, and secondly, you have the capacity to participate in its value. You still have reason to pursue things that do not pass the second test, but their value will not increase your well-being. The same thing goes for the well-being of other people. You will not increase other people's well-being by pursuing value in which those people cannot participate. The view that reasons are all capped by capacities may come of an unholy alliance of two false assumptions, one that reasons are all reasons of well-being, and two, that all reasons of well-being are agent-relative. For further discussion of these and related points see Timothy Macklem, 'Choice and Value', *Legal Theory* 6 (2000) 000, and John Gardner, 'The Mysterious Case of the Reasonable Person', *U Toronto LJ* 51 (2001), 00.

direction. Deliberating about what to do is itself something one does. It is an action (or activity) preliminary to other actions (or activities). As such there are ordinary reasons in favour of and against engaging in it. It follows that deliberating about what to do is not always the best way to do as the reasons that figure in one's deliberation would have one do, for that equation would eliminate the reasons for and against deliberation itself. To put it tersely, there are some values that are badly served by deliberate action but that are well served by spontaneous action.³⁰ Most obviously there is the value of spontaneity itself, which is a constituent of many valuable relationships. But there is also, more prosaically, the value of the time saved by not stopping to deliberate. These values yield reasons against deliberating about whether to perform the actions that one has reason to perform. So the suggestion that the best way to do as reasons would have one do is necessarily to count or weigh up those reasons fails. This failure has nothing to do with exclusionary reasons. Exclusionary reasons are not reasons against deliberating, or against including certain reasons in one's deliberations, but reasons against acting for certain reasons, whether one's action involves deliberation or not.

Much confusion in contemporary moral, political and legal philosophy comes of the assumption that our rational faculties are our deliberative faculties. There is a sense in which, since deliberation is a rational power, action is less rational if not deliberate. But it is only in one respect less rational. In other respects it may be more so.

Somebody who reacts instinctively to save his life does not exhibit all of his rational faculties. Most obviously, he does not

³⁰ This is one interpretation of Bernard Williams' famous 'one thought too many' argument in 'Persons, Character and Morality', Williams, *Moral Luck* (Cambridge: Cambridge University Press 1981) at 18. However another way to read that argument would make it a rendition of our argument in §2 above, to the effect that certain supposed operative premisses are in fact redundant and so not operative at all.

exhibit his capacity for deliberation. But all things considered he is more rational than somebody who exhibits her capacity for deliberation but, thanks to the time wasted in deliberating, dies as she exhibits it. The master capacity that goes under the heading of rationality is the capacity to act for reasons, and this the instinctive self-preserved exhibits par excellence, better than his more deliberative counterpart. After all, he jumps out of the way of the speeding juggernaut while she stays fatally in its path, working out what to do. He is driven by raw fear. This shows the fallacy of the widely touted contrast between reason and emotion. Our fear or anger, so long as it is justified fear or anger, is rationality's friend. One is justified in being afraid or angry just in case one's fear or anger is grounded in facts which are such that if they make one afraid or angry one is thereby more likely to react to them in a measured way than if one measured the way for oneself. Naturally, as we already mentioned in §1, it is possible to perform an unjustified action out of justified fear or anger. Nevertheless the justification of one's fear or anger lies in the likelihood that one will not perform such an action.³¹ Like all justification the justification is rational. If asked why one was fearful or angry one gives one's reasons.

So the issue of deliberation is a distraction. Exclusionary reasons are reasons that regulate the reasons for which one acts.

³¹ Naturally in determining whether the action that one performs out of emotion is justified one must take account of the value of expressing that emotion. One should not assume that an emotion is justified only if it tends towards the performance of actions that would be justified *apart from* the value of expressing that emotion. The test is whether the actions would be justified *taking account of* the value of expressing the emotion. Thus our view of the justification of emotion is not reduced to a purely instrumentalist one. However some emotions are more instrumental than others. The justification of actions expressing grief or guilt, for example, is typically heavily affected by expressive value, whereas fearful or disgusted actions typically take more of their justification from facts other than the fact that they express fear and disgust, making those latter emotions more purely instrumental in their rational role.

The sense that they are paradoxical does not come of the thought that *deliberately* doing what reasons would have us do is rationally superior to doing the same thing *intentionally but without thought*. Rather it comes of the thought that *intentionally* doing what reasons would have us do (whether with or without thought) is rationally superior to *accidentally* doing what reasons would have us do, where this means doing as those reasons would have one do but for some *other* reason. The question raised by the second rationalist myth is whether there is something fishy about accidentally doing as a reason would have one do. Is this not to defy or betray one's own rationality?³²

To see why not one has to understand the true sense in which reasons are there to be followed, and the true sense in which rationality would have one act on the balance of reasons. What rationality strictly speaking asks is that one always act for an *undefeated* reason. The balance metaphor that is in common use focuses attention on cases in which reasons are defeated in ordinary conflicts with ordinary reasons for rival actions. In other words they are outweighed. But reasons that are excluded by exclusionary reasons are also defeated reasons. When rationality asks that one act on the balance of reasons, accordingly, it asks that one act for a reason that has neither been outweighed nor excluded. This formulation reveals that the reality is a kind of *tertium quid* between accident and intention. There is nothing wrong with doing as a reason would have one do but not for that reason (in other words, accidentally doing as the reason would have one do), so long as one does do it for *some* undefeated reason (and in that sense intentionally engages with reasons). This sharpens the test of justification that we proposed in §1. In §1 we said that to show that one's action was justified one must show that reasons did indeed support the course of action that one

³² This is the basis of Robert Paul Wolff's attack on authority in *In Defense of Anarchism* (2nd ed, Berkeley, Ca: University of California Press 1998) ch. 1. Raz replies, in defence of exclusionary reasons and hence in defence of authority, in *The Authority of Law* (Oxford: Clarendon Press 1979) ch. 1.

engaged in and that one acted for one of those supporting reasons. Now we state this biconditional test more technically but also more tersely. To show that one's action was justified is to show that one performed it for an undefeated reason. (And by the same token, to show that one's belief is justified is to show that one held it for an undefeated reason, to show that one's fear is justified is to show that one felt it for an undefeated reason, and so on.)

It may be objected that our talk of defeated reasons illicitly glosses over a major difference between reasons that are outweighed and those that are excluded. When a reason is outweighed, that actually makes a difference to what one ought to do. The correct action is the one supported by the outweighing reasons and vainly resisted by the outweighed ones. On the other hand, the objection goes, an excluded reason is, by our own testimony, merely one for which one ought not to act. The device of exclusion is simply an indirect device for making it the case that one does what one ought to do anyway, given the weighing of ordinary reasons. So understood, exclusionary reasons belong to the philosophy of means rather than ends.³³

But the contrast here is a false one. The fact that a reason is defeated, whether by outweighing or by excluding, is entirely irrelevant to the question of whether it would be better to do as

³³ This contrast between the philosophy of means and the philosophy of ends has dogged the so-called 'rule utilitarian' approach to moral reasoning. The great insight of rule utilitarians such as Mill and Sidgwick was that rules cannot be regarded as mere means to the fulfilment of some independently specified end. What one ought to do given the rule was not necessarily what one ought to do apart from the rule. But deep errors of the kind exposed in the next paragraph led to this insight being underestimated and to rule utilitarianism coming to be regarded as a kind of faint-hearted and untrustworthy variant on act utilitarianism. In reality, of the many things that are wrong with rule utilitarianism, most are owed to act utilitarianism. Rule utilitarianism's one great triumph was its breaking free in the respect just identified from its act utilitarian parentage. The issue is famously explored in John Rawls, 'Two Concepts of Rules', *Philosophical Review* 64 (1955) 3.

that reason would have one do. It would *always* be better to do as every reason would have one do, defeated or otherwise, thanks to the fact that every reason corresponds to a value. In a situation of rational conflict reasons pull one in two or more incompatible directions. Given the incompatibility it is not possible to realize the value in both actions. Tautologously, it is best to act for the best, or in other words, not to perform an action that realizes less value than some other action that one might have performed. But of course it would be better still if by some miraculous change in circumstances, as well as doing what the undefeated reasons would have one do, one also did as the defeated reasons would have one do. This applies as much to outweighed as to excluded reasons. Outweighed reasons do not lose their rational force in the sense of being any the less reasons for action, or in the sense that there is any less appeal than there would otherwise be to doing as they would have one do. The only effect of their being outweighed, like the effect of their being excluded, is to eliminate the rationality of acting for them.

6. THE FORCE OF REASONS

We introduced the category of exclusionary reasons by concentrating on cases in which by acting for a certain reason one risked failing to do as that very reason would have one do. In such cases the role of exclusionary reasons is to control counter-productivity. But not all exclusionary reasons exist to serve this purpose. Some are there to make it more likely that one will do what a certain reason would have one do, not by excluding *that* reason but by excluding some *other* reason, often but not always a conflicting reason. Once at the restaurant it is better not to tailor your choices from the menu according to price, for to do so is to miss out on the spirit of the evening. Likewise it is best not to give one's friends financial advice with a view to one's own profit, since in the process one betrays one's friendship. Examples of this type abound. They differ from cases in which reasons are

merely outweighed. Naturally one may have other reasons to try what happens to be the cheapest item on the menu, or to advise one's friend to make the investment that will also, as it happens, be to one's own profit. These reasons may outweigh the reasons to do the opposite. The point is that when this is the case one has other undefeated reasons to do as the excluded reasons would have one do. But one also has exclusionary reasons not to do it for the excluded reasons, in other words not to do it because of the cheapness or the profit as the case may be.

The classic case in which exclusionary reasons perform this function of fortifying reasons against their opponents is the case of *duty* (also known as obligation). To have a duty to do something is to have a reason to do it that, (i) does not depend for its existence on one's goals at the time, and (ii) is also a reason not to act for certain conflicting reasons. The first feature gives duties their *categorical* character. Categorical reasons are those that are not hostage to the prevailing personal goals of the agent to whom they apply. One's reasons to promise are often dependent on what one wants to achieve, but once one has promised the reason created by the promise does not bend to the changing winds of one's ambitions. But for our purposes it is the second feature that is more important. A reason to do something that is also a reason not to act for certain countervailing reasons is a special kind of reason that is labeled by Raz a *protected* reason. A protected reason is not merely the coincidental conjunction of a reason to act and an exclusionary reason not to act for certain countervailing reasons. The point is that the very same fact that is one's reason to act is also one's reason not to act for certain countervailing reasons. The very fact that one promised is both a reason to do what one promised and a reason not to be moved by the fact that doing it is now more inconvenient than it was when one promised. When a reason has this special protected structure we feel its force as *mandatory* force. We are *required* to do what the reason would have us do. Reasons can be categorical but not mandatory and mandatory but not categorical. Duties are

the special and important case of reasons that are both (i) categorical and (ii) mandatory.

Imagine, for example, that a keen follower of the arts is offered the opportunity to stay in a friend's flat in Edinburgh during the Festival. It is a plum opportunity. Although he lives not far away in the Glasgow suburbs he has rarely had the opportunity to get to the Festival, let alone the chance to stay for several days. There are no strings attached and the keys will be left for him with a neighbour. The case for going strikes him as overwhelming. But alas our culture vulture, call him K, finds that his car won't start on the morning that he is due to set off and the local mechanic says that it will take a week to repair. It looks as if he will have to haul himself into town by bus in order to take the train to Edinburgh. All this hassle has only added to the stress that he was beginning to feel at the prospect of being away from work during a week when, as it turns out, a lot of important clients plan to come in. All things considered the whole idea is becoming more trouble than it is worth. There is no way he will enjoy the Festival under such stressful conditions. His enthusiasm for the arts is waning by the minute. So far so good. K's reasons for seizing the opportunity may now be defeated. But things would be different if he had agreed to look after the flat in his friend's absence and because of this his friend, nervous of leaving his flat empty during Festival season, had not hired a paid house-sitter. In that case K is under an obligation to go to Edinburgh as he said he would. What this means is that he has a reason to go to Edinburgh, viz. that he agreed that he would, and this reason has two special features. Unlike his love of the arts, it is a reason which does not ebb and flow with his changing ambitions and moods and enthusiasms. That is to say, it is a categorical reason. What is more, it is not only a reason to go to Edinburgh but also a reason not to stay at home for the reason that going to Edinburgh would now be more awkward or disruptive than he imagined it would be at the time when he agreed. In other words, the reason is also mandatory.

Let us put the categorical character of duties or obligations on one side for a moment. It affects the conditions under which duties exist rather than the force that they have when they do exist. Our focus will be on their force. The force of duties has two dimensions. To put it another way, duties attack their opponents (the reasons for action with which they conflict) on two levels. First of all they have their ordinary weight as reasons for action. Secondly, they have a special, built-in exclusionary force that protects them against conflicting reasons for action, irrespective of the ordinary weight that either they or their opponents have. It might be thought that the second aspect renders the first redundant. Since duties exclude, what is there for them to outweigh? Since their opponents are already defeated why do they need to be defeated again? The simple answer is that the exclusionary force of duties need not be absolute. Typically, the fact that doing something is one's duty means that one cannot plead certain objections to doing it. Often one cannot plead inconvenience or irritation for example. In K's case, where the obligation is assumed by agreement, K cannot plead the everyday inconveniences and irritations that he had reason to take account of at the time when he made the agreement. Yet other reasons against performance of the duty remain available to be acted upon. For instance, if K's house was burgled on his first day in Edinburgh that might well be an unexcluded reason to return home. And if that other reason is weighty enough it may well justify him in abandoning the Edinburgh flat, thereby breaching his duty to his friend.

At this point everything depends on the weight of the reasons left in play, including the weight of the duty itself, now conceived as an ordinary reason for action. It should not be assumed that this reason wields a particularly great weight. Reasons can be mandatory yet trivial. They can be good at defeating by exclusion but poor at defeating by weight. One has a mandatory but not particularly weighty reason to thank one's hosts at the end of a party. The fact that one was tired at the end

of the evening is no reason not to thank them, for that reason is certainly excluded. Yet a whole host of reasons, ranging from the fact that one's baby-sitter has just made an anxious call to the fact that one's host was deep in conversation, may well be sufficient to justify breach of the duty.

This creates an interesting linguistic ambiguity. In one sense of 'wrong' actions are regarded as wrong if and only if they are unjustified. The fact that there can be justified breaches of duty introduces, however, a different nuance of the word wrong. There can be actions that are wrong although justified and that create a special pressure for justification precisely because they are wrong. They are things which, all things considered, it was right for one to do and yet in doing them one did something wrong. In the latter inflection 'wrong' just means 'in violation of a requirement', leaving the question of justification open.³⁴

It is often thought that wrongdoing in this latter sense is of special importance because of the traces it leaves on one's life. The point cannot be denied and yet it can be overstated. Every time one does not do what *any* reason would have one do, be that reason mandatory or otherwise, a trace is left on one's life. This is always in principle a matter of regret, for as we saw reasons do not lose their force as reasons merely because they are defeated.³⁵ Where ordinary reasons are concerned, however, this

³⁴ Some people like to follow W.D. Ross, *The Right and the Good* (Oxford: Clarendon Press 1930) in deploying the label 'prima facie wrong' for actions that are wrong in the latter sense. Similarly they may speak of prima facie duties, prima facie reasons, and so on. This usage is misleading. Prima facie reasons are just reasons, prima facie duties are just duties, and correspondingly, prima facie wrongs are just wrongs. For trenchant criticism of Ross' expression see John Searle, 'Prima Facie Obligations', in Joseph Raz (ed), *Practical Reasoning* (Oxford: Oxford University Press 1978).

³⁵ Bernard Williams' long flirtation with the view that reasons are 'internal' rather than 'external', leading him away from the position we adopted in §§1-3 above, came primarily of the thought that the position we adopted leaves no room for regret. We hope to have demonstrated the opposite. See Williams,

enduring force, like the original force of the reasons in question, is merely *advisory*. It may weigh heavily, but all the many and well known reasons not to regret also weigh heavily and often justify one in getting on with one's life regardless.

Where one fails to do as a mandatory reason would have one do, however, the enduring force of the reason is different. It continues to be a mandatory reason even after one's failure. Just as some reasons not to perform the required action were excluded by virtue of the fact that the action was required, so some reasons to forget all about one's non-performance of the action afterwards are excluded. In other words, it is not so easy to brush aside the failure to do as a mandatory reason would have one do. It cannot just be a matter of pointing out that one would have a better life if one stopped crying over spilt milk. This is true *a fortiori* where the mandatory reason is also categorical, for here changes in one's goals do nothing to eliminate the reason's application, and this is true no less when the time for performance is past. One cannot escape the blemish on one's life by a simple change of direction. The rational pressure that this creates is the familiar pressure for various remedial and purgatory reactions, such as apology, payment of compensation, penance or punishment.³⁶ The appropriateness of each of these reactions is, of course, subject to further conditions. Some, like punishment, may be unjustified if the action to which they are responses is a justified action. But be that as it may, all require violation of a duty.

We just spoke of ordinary (unprotected) reasons for action as having a merely advisory force as opposed to a mandatory force. Notice that this is not the same as saying that they are permissive.

'Ethical Consistency', in Williams, *Problems of the Self* (Cambridge: Cambridge University Press 1973).

³⁶ The same forces create the rational pressure for retrospective emotions such as guilt, shame and regret. The moral importance of these emotions demonstrates the importance of the point made about the expressive (as well as instrumental) value of emotion in note 31 above.

To be sure, there are some reasons which are such that it is permissible not to do what they would have one do, and supererogatory (or ‘beyond the call of duty’) to act for them. These represent another special type of reason involving exclusionary elements. They are not merely unprotected reasons but positively *exposed* reasons. We cannot explore their logical structure here, which is more complex still than that of a protected reason.³⁷ But it is worth mentioning them simply to warn against the familiar tendency to think that reason has only two voices, namely the raised voice of requirement and the conspiratorial whisper of permission. This legalistic contrast unfortunately misses out rationality’s normal speaking voice. Normally reasons are simply our counsel and guide. All special exclusionary components apart, a reason merely supports the action (or belief, emotion, etc.) that it supports. In the absence of any exclusionary force, a reason to perform that action makes that action advisable. If there are conflicting reasons, the fact that one action is supported by weightier reasons than any of the others simply makes it the most advisable action. Perhaps I do a better job of my shopping if I go early on a Saturday, ahead of the crowds, rather than later in the day. And perhaps all else is equal: I am not inclined to have a lie-in, I have nothing else planned for either part of the day, and so on. If I go later I clearly do the less advisable thing. Rationally speaking, I err. I do the wrong thing in the first sense we mentioned. But I do nothing wrong in the second sense. It is not mandatory for me to go shopping at the most advisable time. For, so far as the story goes, I have no protected reason to do so.

A common assumption, we suspect, is that in cases like this a protected reason automatically swings into action at the final hurdle. Our conclusion, clearly, is that it is not rational for me to defer my shopping. So surely going shopping early is my only

³⁷ See Raz, ‘Permission and Supererogation’ in *American Philosophical Quarterly* 12 (1975) 161.

option *except on pain of irrationality*? I may be under no requirement to go shopping now if my reason to do so is defeated, but since it is undefeated and every conflicting reason is defeated, doesn't the general requirement to be rational, the general requirement to act for an undefeated reason, swing in and turn this into a mandatory reason at the end?

We doubt whether there is any such general requirement to be rational. Some people in some special roles, such as judges and other bureaucratic officials, are under duties to be rational in those roles. Moreover we all do have various duties connected with respect for our own rationality and that of others. Yet it is a long way from this to the conclusion that rationality itself is required of us across the board. But even if it is, notice, this claim confirms rather than denies the most important point we made in this section. To understand the idea of a general requirement to be rational one would need to understand what a requirement is that distinguishes it from an ordinary reason for action. It is tautologous to claim that we have reasons to be rational. But it is not tautologous to claim that we are subject to a *requirement* to be rational, or (what is more) that this is our *duty*. Since these claims are not tautologous there must be something more to requirements and duties than there is to ordinary reasons. This section explained what that 'something more' is. Requirements (or mandatory reasons) are protected reasons, meaning that they are reasons for action that are also reasons not to act for certain reasons that militate against that action. Duties (or obligations) are exactly the same, but with the extra feature that they are also categorical, meaning that they are not hostage to the prevailing personal goals of those who are subject to them.

One explanation of why people have been anxious about conceding the fundamental agent-neutrality of reasons has been the thought that if one person's reasons to act automatically yield reasons for everyone else to contribute to the first person's acting as those reasons would have her act, then our horizons are filled

with extremely demanding duties.³⁸ After all, the thought goes, reasons are mandatory, and reasons to serve the goals of others depend on *their* goals, not on our own goals, and so appear to be categorical. Is there any room left, the question arises again, for me to have goals of my own?

We provided a partial response to this anxiety in §4. We said that there are limits to the extent to which it is possible for one person to contribute to another's pursuit of her goals, and that possibly, in view of the likelihood of over-enthusiastic intervention, people should adopt, within limits, an agent-relative approach to life. Now we add another point. Even if one could do a lot to help another achieve her goals it does not follow that one has a *duty* to do so. In fact there is no reason to think that even if another has a duty, whatever one can do to help them perform their duty is one's own duty.

The fact that one person's having a reason automatically gives all people reasons to contribute to the first person's doing as the reason would have him do does not entail that, if the first person's reason is a categorical mandatory reason, the reasons of all others involved are categorical and mandatory as well.³⁹ To establish that they are, one would need to show that relative to each person there are reasons not to act on certain countervailing reasons, including those that relate to that person's goals. This is very often a tall order. *A fortiori* it is a tall order to show that one has duties to contribute to every one else's fulfillment of their goals. Naturally this does not mean that one never has such duties. It just means that the existence of all such duties depends

³⁸ For powerful expression and critique of this anxiety see Samuel Scheffler, *Human Morality* (New York: Oxford University Press 1992).

³⁹ Technically: One has reason to do whatever is sufficient to contribute to the satisfaction of any (other) reason, whether or not it be necessary. However one does not necessarily have a duty to do whatever is sufficient to contribute to the performance of any (other) duty. On these aspects of the logic of practical reasoning see Anthony Kenny, *Will, Freedom and Power* (Oxford: Blackwell 1975) ch. 5. Throughout this paper we have broadly sided with Kenny's conclusions.

on the arguments that support the existence of that particular duty, in other words, that support one person's contributing to the pursuit of another person's goals independently of her own and irrespective of at least some of the price that has to be paid.

7. THE ROLE OF CHOICE

So in two ways, we have discovered so far, rationality hems us in less than some have been accustomed to think. But so far we have neglected a third, and some may say more important, way in which rationality leaves our options open. We have said nothing of the role of incommensurability. Leaving aside for a moment the complicating effect of exclusionary reasons, two alternative actions are incommensurable if neither is supported by weightier reasons than the other and yet the reasons on the two sides are not equal in weight either.⁴⁰

Suppose that a young woman, call her L, has two alternative activities before her this afternoon. She can go to a matinee screening of *Casablanca* with her aged father or she can have a game of tennis with her brother. Both are valuable pursuits. They will make different members of her family happy in different ways, will provide her with different kinds of relaxation, and will each contribute differently to her pursuit of her own goals, for she is both a film buff and an avid tennis player. Neither outing can be postponed, for *Casablanca* is being screened for one afternoon only and her brother can only get this afternoon off from work. What is she to do? It is not as if one alternative will bring greater pleasure than the other, or will make more interesting demands than the other, or will meet anyone's needs more than the other. There seems to be nothing to choose between them. L may conclude that the two alternatives are equally supported by reasons.

⁴⁰ Raz, *The Morality of Freedom*, above note 24, at 322

But this is not the only possible conclusion. If they are equally supported by reasons then any little extra reason on one side or the other will make it the case that that alternative is rationally superior. Suppose that L's father rings to tell her that *Casablanca* is to be screened in a new print. Assuming that the quality of the new print strengthens her reasons to see the film, albeit ever so slightly, L should presumably regard her father's call as settling the matter. She was previously unaware of one of the reasons that militated in favour of the film, and in her ignorance of that reason had calculated that all the other reasons were evenly balanced. Assuming she does not now regard that calculation as having been mistaken in any other way, she now knows that it was mistaken in one way, namely that one reason was omitted. If without that reason the two alternatives were equal in respect of rational support they cannot possibly be equal with that reason added.⁴¹ Yet L may conclude that there is still nothing to choose between the alternatives. If she is right so to conclude, both before and after she is furnished with the extra information, the case is one of incommensurability. The reasons to go to the film with her father neither outweigh by nor are outweighed by the reasons to play tennis with her brother. And yet they are not of equal weight either.

Reasons of equal weight are of little day-to-day importance in practical reasoning, for the necessary small adjustment to eliminate the equality almost always lies close at hand. Typically one can strengthen one's reasons to perform one action rather than the other simply by adjusting one's goals, so adjusting the value of personal engagement, which promptly tips the balance. At this point the reasons on the other side are defeated by weight. But things are different where the reasons on the two sides are not equal but incommensurable in weight. In cases of incommensurability adjustments to the reasons on one side of the

⁴¹ This point (relied upon by Raz) is owed to John Mackie, 'The Third Theory of Law', *Phil & Pub Aff* 7 (1978) 1, at 9.

argument do not necessarily lead to reason favouring that side. This means that incommensurability of reasons has far more day to day importance than equality of reasons, for the small adjustments that one can call upon to resolve the latter do nothing to resolve the former. That being so, the principle of rationality that commends action for an undefeated reason does not adjudicate as between the two alternatives. Still assuming no exclusionary complications, both actions are supported by undefeated reasons. Yet there is no point in trying to tip the balance for there is no balance to tip.

Some people offer what they take to be an alternative interpretation of L's predicament. They say that if the new, admittedly rationally relevant information (that the print of Casablanca is a fresh one) fails to tip the balance in favour of the film, then that shows that the two alternatives before L, while possibly not equal, were and remain *roughly* equal.⁴² But what does this mean? We think that it probably means that L is somewhat uncertain of the weight to be attached to some or all of the various reasons involved, so that her conclusion before her father calls (that the reasons were equally balanced) was necessarily an approximation.⁴³ Yet, the story must continue, the additional reason provided by her father does not exhaust the margin of error in the original approximation, and therefore provides no reason for her to revise it. The alternatives remain roughly equal.

It may be that there are some cases like this.⁴⁴ But if so they are different from cases of incommensurability. The notion of rough equality belongs to the realm of indirect reasoning strategies, for it is a strategy (one may doubt whether it is a good

⁴² For example, Derek Parfit, *Reasons and Persons*, above note 4, 430-32; James Griffin, *Well-Being* (Oxford: Clarendon Press 1986) 80-81.

⁴³ For another possible explanation of what some writers may mean by rough equality, see Macklem, 'Choice and Value', above note 29, where the position outlined in the text is also further explored.

⁴⁴ We say 'it may be' because we disagree with each other here.

one) designed to cope with uncertainty about the weight of reasons. Given the uncertainty, so the story goes, one has an exclusionary reason not to act for very small additional reasons in favour of one side. But in cases of incommensurability there is no reason not to act for the very small additional reasons that may be brought to bear in the vain hope of resolving the conflict by rational means. One may act for these reasons to one's heart's content but one will still be no closer to a decisive rational direction. The rival paths will still both be supported by undefeated reasons and it will still be impeccably rational to take either.

Although we are both believers in the pervasive role of incommensurabilities in rational life this is not the place to defend that belief. However it is important to appreciate its significance. One may ask why the thought of incommensurability makes people so edgy. Even among those who have a duty to be rational, or in other words, a duty to act for an undefeated reason, it is not clear why any should be thought to have a duty, or even a reason, never to act *against* an undefeated reason.⁴⁵ Perhaps the feeling that people have such a reason is a further reflection of the perniciousness of the rationalist myths that we referred to in §5. These myths were the work of those who read into rationality a promise of security against chance, luck, arbitrariness. The will is regarded as an arbitrary factor in need of rational control. But the will is not arbitrary in any worrying sense so long as it always responds to undefeated reasons. This condition is amply met in normal cases of incommensurability in which reasons on both sides are undefeated. Rationality led one as far as it promised to do by

⁴⁵ Much work in jurisprudence, asserting or denying the existence of 'right answers in hard cases' has been based on the assumption that judges do have such a reason, or even such a duty. Occasionally public expectation may generate such a reason, but by and large, in our view, judges have nothing to fear in the fact that they did not defeat the reasons against their decision, so long as they gave undefeated reasons for their decision.

leading one to this situation. It promised no more. It was merely a bizarre unfounded expectation of those frightened by the human condition that it would lead one further and tell one which way to choose given that both choices were already *ex hypothesi* rationally defensible.

It could be replied that the role of choice is already amply accommodated by the considerations we adduced earlier. We already made space for people to have their different goals and for this to affect what they ought to do. But the role of will that we introduce here is more radical. It explains not only how we bring our goals to bear on what we ought to do but how we are ever set free enough to acquire those goals in the first place. So far as the earlier arguments were concerned our goals exerted rational pressure on us, justifying some courses of action that would not have been justified without them and correspondingly withdrawing justification from others. But in the normal situations of incommensurability mentioned in this section both alternatives are *ex hypothesi* justified, and one's choices do not tip the rational balance. Rather they send one in one direction rather than the other without making it a matter of *rational* determination that one went that way.

We just spoke of 'normal situations of incommensurability' as opening up our options. This was to leave space for a special class of cases in which the effect of incommensurability is, disturbingly, reversed. Rather than offering us more than one justified way forward, incommensurability here denies us any justified way forward and relegates us to the situation in which the best we can hope for is an excuse. The situation we have in mind is a special case in which the conflict of incommensurables is a conflict between two or more *protected* reasons. The basic problem is that the reasons are protected against each other. Each excludes the other from consideration. When this happens the normal effect is that the conflict is decided by the weight of the reasons *qua* exclusionary. The question is: which of the two reasons not to act for a reason is the weightier? But suppose the

two exclusionary reasons are themselves incommensurable. In that case neither defeats the other. Thus neither prevents the other from exerting its exclusionary force. Thus both of the reasons for action are defeated by being excluded. Thus there is no way to act for an undefeated reason. This is the core case of what is colloquially known as a moral dilemma. Many things referred to as moral dilemmas do not quite reach this pitch of tragedy. But all carry intimations of it. All remind us that the supposed security of rationality is not only compromised in the direction of giving us all elbow room to follow our wills, but occasionally is also compromised in the direction of locking us into a situation in which, whatever our virtues, we are doomed to act against the basic maxims of rationality.⁴⁶

When the concentration camp guard offered Sophie the choice between sending both of her children to their deaths and choosing which child to save, his scheme was to imperil her rationality in just the way that we have described. He hoped that by facing her with two conflicting and mutually exclusionary incommensurable duties towards her children, namely, the duty to protect them and the duty not to display more love for one than for the other, he would leave her in the situation of having no justified alternative before her, and hence with her rationality demeaned (in addition to all the other horrors that she will experience). Most readers of *Sophie's Choice* believe that the guard failed in his plan. Sophie was justified in saving one child while condemning the other to death. Either this was not a display of more love for the saved child, most readers think, or the duty not to display more love for one child than the other was outweighed by the duty to protect. Of course this leaves Sophie in the rational predicament that we described in §6. She still was in breach of one of her duties and the scars of this remain

⁴⁶ We suspect that these are the cases that Williams means to label 'tragic' in 'Conflicts of Values', in Williams, *Moral Luck*, above note 30, at 74, although his remarks are open to more than one interpretation.

with her.⁴⁷ You may say that this is bad enough. But it falls short of the moral destruction that the guard aspired to. He aimed not only to damage Sophie as a parent but to damage her as a rational human being by putting her in a situation in which nothing she could do would be justified.

8. THE STANDARD OF REASONABLENESS

A number of contemporary moral philosophers distinguish rationality from something else called reasonableness.⁴⁸ As they describe it, the standard of reasonableness is more accommodating than rationality, for it leaves latitude for differences of opinion, for rival views of what it is that rationality demands. It must be said that by and large it is obscure in these views what it is about rationality that the more accommodating test of reasonableness is needed in order to escape. All too often it seems that these views rely on a parody of rationality, in which the rational person is one who has developed just one aspect of her rational functioning at the expense of all others. On this picture the rational person is a single-minded deliberator about everything, or a pure instrumentalist who uncritically takes the value of her goals for granted and pursues them ruthlessly, or a high-minded person of principle who scrupulously abandons all partiality. We agree that it is possible to use the word rational to pick out these narrow and limited abilities, but the abilities so picked out are indeed narrow and limited and the mastery of one at the expense of the others is not rationality in any sense in which that could constitute a valid complete ideal for human

⁴⁷ Some hold that the very presence of such scars, if they are rational, must reflect the presence of some incommensurability. For excellent discussion which brings out both the element of truth and the element of falsity in this view, see Jonathan Dancy, *Moral Reasons*, (Oxford: Blackwell 1993) 120ff.

⁴⁸ John Rawls, 'Kantian Constructivism in Moral Theory', in *Journal of Philosophy* 77 (1980) 515; Scanlon, *What We Owe to Each Other*, above note 13, 22-33.

agents. Properly understood, rationality is just what we have described in this paper. It is exactly the same as reasonableness. As we have urged, it is simply the capacity and propensity to act (think, feel, etc.) only and always for undefeated reasons. As an ideal the reasonable person, or the rational person (as she might just as well be called) is the person who fully exemplifies this capacity and propensity, and does so in her beliefs, emotions, attitudes, actions and so on. Wherever there are reasons, she does not defy their force.

On this view 'reasonable' means much the same as 'justified'. A reasonable action is a justified action, a reasonable belief is a justified belief, a reasonable emotion is a justified emotion, and so on. By the same token, a reasonable person is no less and no more than a person who is justified in whatever aspects of her life call for justification. In the law, the standard of the reasonable person is sometimes used to test beliefs, and the test of reasonableness is then the test of justification. It is true, as we said in §1, that acting on the strength of a justified belief is not the same thing as acting justifiably where that belief is false. Such an action is not justified but excused. It may be that the idea that an excused action is nevertheless the action of a reasonable person is what has given credence to the idea that a standard of reasonableness offers greater latitude than the standard of rationality. It may be that this has led some to conclude that rationality is what justifies while reasonableness excuses. Yet this is a mistake. The only thing that can excuse is a justification. That is to say, the justification of one human response, for example, a belief or emotion, is what excuses another human response, normally an action. The standard of the reasonable person is one of justification throughout. It is the standard of rationality, a standard that one meets if one has undefeated reasons and responds to those reasons, be it in one's beliefs, or emotions, or attitudes, or actions, as the case may be.