please cite published version:

in Daniel Star (ed.), *The Oxford Handbook of Reasons and Normativity*(Oxford: Oxford University Press, 2018), pp. 423-42.

**Weighing Reasons**

Garrett Cullity

Normative reasons for action are reasons to do things – reasons *for* doing them. We think about these when deliberating about what to do. Often, there are reasons for and against different options, and some are stronger than others. In deliberating, we are trying to figure out their comparative strengths, or weights, and reach a conclusion about which action is most strongly supported overall. This chapter is about what is involved in doing that.

 We can start with the following Simple Picture. As we ordinarily talk and think about them, normative reasons are facts ([Raz 1999: 17](#_ENREF_24)). Ordinarily, reason-reports either take a propositional form, or use shorthand expressions that can be converted into that form easily enough. The reason for cancelling the picnic is that it is forecast to rain; when you say that you cancelled it “because of the weather”, you are understood as referring to that fact. A fact is a reason for action when it bears a special, normative relation to it – the relation of normative support, or “counting in favour” ([Scanlon 1998: 17](#_ENREF_29)). This relation is a vector, which is to say that it has a direction and a magnitude ([compare Broome 2013: Sect. 4.3](#_ENREF_4)). There are only two directions a reason for action can have: either for or against.[[1]](#endnote-1) But the magnitude has no such restriction; it can vary just as much as physical weights can. The weights of the reasons for and against an action can be summed, and this determines how strongly supported the action is, overall, by the reasons that bear on it, for and against. The reasons for doing something are sufficient when no other action is more strongly supported, overall, and decisive when this action is most strongly supported ([Parfit 2011: 32-3](#_ENREF_23)) – then you ought to do it. You ought to perform the action with the greatest net balance of reasons in favour over reasons against: the one there is most reason to perform ([Setiya 2007: 13](#_ENREF_33)). Rational thought and action then correctly reflects this balance. Rational deliberation is weighing reasons correctly. Acting rationally is doing what you have sufficient reasons to do ([compare McNaughton and Rawling 2004: 114](#_ENREF_20)).

 This view of what is involved in weighing reasons seems adequate in many cases. In deciding which holiday to take, you are apparently weighing your prospective enjoyment of each of the alternatives against its cost, and trying to ascertain which offers the best value for money: that is the one that you ought to take, and the one it is rational to choose. But in what follows, we are going to explore the ways in which the Simple Picture appears to require modification and refinement. The first four sections examine some of the ways in which talk of *the* weight of a reason may need improvement. Sections V and VI look more closely at the relationship between reasons and rationality. Sections VII and VIII ask whether there are ways in which a reason can be defeated which are not kinds of outweighing. I shall conclude that while in some respects the Simple Picture does need to be corrected, in others the jury is out.

**I: Derivative Reasons**

According to the Simple Picture, you ought to do what there is most reason to do. This needs clarification, since there are different ways of using the word “ought”. Suppose my pregnant wife tricks me by convincingly pretending to go into labour. Ought I to take her to hospital? In one sense Yes: if she does appear to be going into labour, I behave defectively if I fail to take her, and this is naturally expressed by saying that I ought to have taken her. But in another sense No: since taking her to hospital will actually waste everyone’s time, we could just as naturally express *that* by saying that it is not what I really ought to do. We therefore need a distinction. The second “ought” is objective: it is determined by the facts that do actually count for or against my actions independently of whether I recognize it. The first is subjective. It depends on my beliefs about the objective reasons – or perhaps, the evidence available to me about those reasons, or my justified beliefs about them.... In fact, there seem to be several different subjective “oughts” ([Parfit 2011: Ch. 7](#_ENREF_23)). But let us set them aside here. What we are primarily interested in when we deliberate is figuring out which action really is supported by the objective reasons.[[2]](#endnote-2) So let us concentrate on that.

 We can next ask which of our objective reasons carry independent weight. Ordinary discourse is a treacherous guide to this. If my wife *has* gone into labour, it would often make sense to cite that fact as the reason for taking her to hospital. It also sounds true to cite as a reason for taking her the fact that the hospital is a safe environment for giving birth, or that she is relying on me for transport. But it seems wrong to count these as separate reasons, and add their weights to each other when assessing how much normative support there is for my action overall.

 It is tempting to legislate that there is really only one reason here. The reason for taking my wife to hospital, properly stated, is that it is the most convenient means of ensuring that she gives birth in a safe environment.[[3]](#endnote-3) But saying that implies that most ordinary reasons-talk is false, and that seems unnecessarily drastic. Instead, we can take our cue from causal-explanatory discourse. Talk of “the cause” of an event usually picks out one salient feature of a larger causal mechanism, with the choice being determined by the pragmatics of the conversational context ([Lewis 1986](#_ENREF_19)). Likewise, in normative discourse, pragmatic factors affect what we identify as “the reason” for an action, but that does not make ordinary reasons-discourse false. Rather, it is made true by the relation that the salient feature (here, that my wife has gone into labour) bears to the larger normative-weight-bearing fact.

 Should we say, then, that the fact that my wife has gone into labour is a reason with zero weight? No: that is unnecessary too. Instead, we can call it a *derivative* reason. It derives its reason-giving weight from the fact that taking her to hospital is the most convenient means of ensuring that she gives birth in a safe environment. The first fact has reason-giving weight, but its weight cannot be added to that of the reason from which it derives.

 Saying this is coherent. After all, we should make an analogous claim about the physical weights of parts and wholes. My arm has a weight, but we make a mistake if we add it to the weight of my whole body in calculating how hard the scales will be pushed down. However, the disanalogy is worth noticing too. The relation between the two practical reasons in our example is not that of part and whole. And while we might well think that the physical weight of a whole derives from the weights of its parts – it is the sum of those component weights – the derivation-relation for reasons runs the other way: the “smaller” reason derives from the “larger” one. An intriguing question is just what kind of relation normative derivation is, if it is not mereological. But let us not pursue that. The important point here is that we need a first refinement to the Simple Picture. Although derivative reasons have weights, we cannot add them to the weights of the reasons from which they derive, nor to other reasons that derive from the same source.

**II: Holism**

When we consider the contribution that the weights of components make to the overall weight of some larger whole, “holism” might be used as a name for two different theses.[[4]](#endnote-4) One is a thesis of non-compositionality: the overall weight of the whole is not always the sum of the weights of its parts. This is the analogue of Moore’s doctrine of “organic unities” concerning the relationship between the value of a whole and the values of its parts ([1903: Sect. 18-20](#_ENREF_22)). The other is a thesis of circumstance-relativity: the weight of a weight-bearer can vary depending on the rest of the whole set of circumstances by which it is surrounded. It is this second thesis that is now usually meant by talk of “holism” in the theory of reasons.[[5]](#endnote-5)

 The difference between the two kinds of “holism” is illustrated by physical weight. This is compositional: the physical weight of a whole is the sum of the weights of its parts. But it is circumstance-relative, on both of its main definitions. On the gravitational definition (as the force exerted on a body by gravity), objects weigh less out in space than they do on the surface of the Earth. On the operational definition (as the force exerted by an object on its support), you weigh slightly less when the Moon is overhead.

 Both kinds of holism are true of derivative reasons. We saw their non-compositionality above: the weights of reasons deriving from the same source cannot be summed. But when a reason is derivative, it is also obviously circumstance-relative. If the fact that my wife has gone into labour derives its status as a reason for taking her to hospital from further facts about safety, then changing those further facts will change its status as a reason. In circumstances in which the hospital is unsafe, the fact that she has gone into labour ceases to be a reason for taking her there.

 So the Simple Picture needs another refinement. When it talks of “the weight” of a reason, it needs to allow that the assignment of weights to derivative reasons is circumstance-relative. It would be clearer to talk of the weight of a-reason-in-a-set-of-circumstances.

 But what about non-derivative reasons? They might still be atomistic, not holistic. Indeed, the most obvious way of explaining the circumstance-relativity of derivative reasons encourages that view. In the example just given, we explain the circumstance-relativity of the derivative reason (the fact that my wife has gone into labour) by pointing out its dependence on whether the further facts from which it derives its status as a reason are present or not. If all explanations of circumstance-relativity worked that way, then non-derivative reasons could well be atomistic. Moreover, only non-derivative reasons are relevant to any judgement about the comparative weights of reasons that supports a conclusion about what there is most reason to do overall. Derivative reasons are irrelevant to that, since they add no further support to the actions for which they are reasons, beyond the reasons from which they derive. If non-derivative reasons are atomistic, the holism of derivative reasons becomes a relatively superficial phenomenon.

 Broadly speaking, there are two main options in defending atomism about non-derivative reasons. One is to hold that there is ultimately only one non-derivative reason for action. The usual candidate is happiness: one’s own, according to the classical egoists; everyone’s, according to the classical utilitarians. Such views can allow that we have many different reasons, but they all derive that status from their contribution to happiness. Since that is the only non-derivative reason for action, it must be the sole determinant of what we ought to do, in all circumstances. Monism about non-derivative reasons entails atomism about them.

 Alternatively, one might hold that there is more than one non-derivative reason, and that they all behave atomistically. The ethical theory of W.D. Ross has this form. Ross recognizes fundamental, non-derivative ethical reasons of several distinct types – “prima facie duties”, he calls them. The overall rightness of an action – one’s “duty proper” – is seen as depending on the presence of these reasons, and if they conflict, how comparatively strong they are.[[6]](#endnote-6) And any instance of one of these reasons carries the same weight in favour of an action, in all circumstances.

 Ross’s theory seeks only to identify the determinants of duty, rather than all non-derivative reasons for action. But its attraction comes from inheriting the clarity of the Simple Picture: Ross’s view is the attempt to interpret the fundamental structure of ethics in terms of that picture. This gives us a way of thinking about ethics combinatorially, explaining its complexity in terms of the interaction of a relatively small number of simpler fundamental elements: the atomistic, non-derivative ethical reasons from which others derive.

 Theorizing of this Rossian form cannot be swiftly dismissed on the grounds that practical reasons are obviously holistic. The most obvious examples of holistic reasons are derivative ones, and explanations of those tend to *support* the search for atomistic non-derivative reasons, as we just saw. Moreover, if a convincing case could be made that the items on Ross’s own list of “prima facie” duties fail to provide us with atomistic non-derivative reasons, that might only show that he has offered the wrong candidates. The only way to resolve this issue seems to be by subjecting the Rossian framework to the test of explanatory adequacy: examining its ability to give a cogent explanation of what strikes us as practically important, by considering a range of examples in detail. The most convincing case for atomism or holism about non-derivative reasons would be to show that one of them does better, overall, in supplying illuminating explanations of the ways in which reasons bear on actions of different kinds. We return to this topic later, in considering whether reasons can be subject to “undermining”.

**III: Incommensurability**

“Weighs more than” is a transitive relation. That A weighs more than B and B weighs more than C entails that A weighs more than C. For physical weights, it is also true that if A weighs more than B and B does not weigh less than C, then A weighs more than C. However, when we assign weights to reasons, there are cases in which the analogue of the second claim apparently fails. Indeed, such cases seem very common. In choosing between restaurants, I can have a stronger reason to go to Angela’s than Bernie’s, but at the same time it can also be true that I do not have a stronger reason to go to the Chopin concert instead of Bernie’s, and nor do I have a stronger reason to go to Angela’s instead of the Chopin concert. If so, how do I compare the strengths of the reasons to go to Angela’s versus the Chopin concert? The first is not stronger; but nor can it be weaker, or equally strong – that would violate the transitivity of “less than” and “equal to”. So we are forced to say that neither reason is stronger, and nor are they equally strong.

 This is usually described as the phenomenon of “incommensurability”.[[7]](#endnote-7) What does it imply about the weights of reasons? Talk of “incommensurability” suggests the view that goods of some kinds (such as pleasant meals and fine concerts) are so different that the strengths of the reasons we have to pursue them cannot be measured on the same scale. But that goes too far ([Griffin 1986: Ch. 5](#_ENREF_13)). If the concert will be a once-in-a-lifetime opportunity to hear a great performer, you could surely have more reason to hear it than to go out to dinner. The failures of commensurateness are local, not global. Moreover, there are features of our thought about reasons that still seem to require thinking of their weight not just ordinally but cardinally, as measured by a ratio scale. When accepting an invitation will cost you two wasted days and not just one, the reason to decline it is twice as strong.

 What the incommensurate cases show is that such comparisons are subject to local breakdowns. These breakdowns appear not to be merely due to our epistemic limitations. I could have as much information as I liked about the two restaurants and the concert and still be left with the same set of comparative judgements. And that would leave me unable even to make an ordinal comparison of the relative strengths of the reasons to dine at Angela’s or go to the concert, let alone to assign to them a cardinal magnitude on a ratio scale.

 This carries two main implications for the Simple Picture. The first is that it needs to abandon the thought that a reason is a vector, with a cardinal magnitude. In cases like the one just described, no unique numerical value can be assigned to the incommensurate reasons.[[8]](#endnote-8) The second is that a corresponding qualification must be made to the talk of “summing” the weights of reasons to determine which action is most strongly supported overall. When an action is supported by one reason, the addition of another (provided it is not derived from the same source) makes the overall support for the action stronger. So the weights of reasons can in this sense be added to each other. But given the local breakdowns we find in our ability to make such comparisons, summing of the weights of reasons cannot be thought of as the addition of numerical values.

**IV: Permission and Requirement**

Next, we can examine the Simple Picture’s claim that

The reasons for doing something are sufficient when no other action is more strongly supported, overall, and decisive when this action is most strongly supported – then you ought to do it.

This claim, too, is contested by some philosophers. To see why, consider this pair of cases:

Case One. I am looking for a relaxing way to spend my afternoon. Two options are to go for a beautiful bike ride in the spring sunshine, or to lie in a hammock at home reading poetry. Although the bike ride would be more enjoyable, I decide to stay put and read instead.

Case Two. As someone new to university teaching, I am finding it stressful. I worry about my classes, and feel that I don’t measure up to my expectations of how good I should be at my job. On some days, I feel like resigning to escape the pressure, but I know that the feeling will pass and this is a job that will suit me in the long term. Despite that, I go ahead and resign.

These cases illustrate two different ways in which we can experience the force of the reasons we have for doing things. In Case One, the most the reasons seem to do is to render either action eligible for choice ([Raz 2006: 109](#_ENREF_26)). There is a sufficient reason for either action, although one of them is more strongly supported. The fact that one action is more eligible doesn’t stop the other from being eligible too. There is a sense in which I am permitted to do either. But in Case Two, things are different. Here, the reasons make one of the actions mandatory. There is a decisive reason not to wreck my career, which defeats the reason to escape my current discomfort. Here, one action is required and the other ruled out. So Case Two fits the Simple Picture; but Case One does not.

 If the Simple Picture is faulted on the strength of examples such as this, what should we replace it with? One suggestion is that we should distinguish two different kinds of conclusions about the bearing normative reasons have on actions: conclusions about which action is best, and about what you ought to do. Reasons of enjoyment, according to this proposal, only ever support conclusions of the first kind, not the second. It may be best to go on the bike ride, but it is not true that I ought to: reasons of enjoyment are not enough to support that kind of practical conclusion. They “entice” by making actions attractive, but do not require them ([Dancy 2006](#_ENREF_9)).

 But that suggestion faces two problems. There is a problem with the idea that the fact-type to which a reason belongs determines the type of conclusion it is capable of supporting. In Case Two, the compromise to my career is a requiring reason. But suppose we change the example, and weaken the reason: I could move to a more agreeable city, but it would involve some career compromise. Now the role of the reasons is to make the options attractive, without requiring: they apparently permit me to resolve the issue either way. Similarly, by strengthening reasons of enjoyment they could do more than just entice. If you win tickets for a round-the-world holiday that you will immensely enjoy, but fail to go because you are too lazy to collect the tickets, surely we can correctly say that you ought to have gone.

 In response, one might retain the distinction between two types of practical conclusion while dropping the claim that the fact-type to which a reason belongs determines which of these conclusions it is capable of supporting. However, this still faces a second problem: the following question needs an answer. If normative support – the support that reasons give to actions – is just one relation, how does it lead to two different conclusions? Why is it not the case that you ought to do what is best? It is natural to answer this by appealing to the margin by which the best action is more strongly supported than the alternatives. For the balance of reasons to support an action decisively, and thereby make it into what you ought to do, it is not enough that there is a greater net balance of reasons supporting it over the alternatives: there must be a much greater net balance. In Case One, this margin is relatively small: while going is better than staying, it is not so much better that I ought to do it. In Case Two, the reasons to preserve my career are much stronger than the reasons to resign. The judgements in the examples of moving to a more agreeable city and the round-the-world holiday can be explained in the same way. However, when we turn to other cases, this suggestion fails. An obvious difference in the strength of reasons can be decisive, even when it is small. If you know that the same computer can be bought for $20 less from one of two equally reliable retailers, you ought to do that, and it would be irrational not to.

 A different line explanation is offered by Joshua Gert ([2007](#_ENREF_11)). Reasons for action have two different kinds of strength: justifying strength and requiring strength. A requiring reason is one capable of making irrational an action that would otherwise be rationally permissible; a justifying reason is one capable of making rationally permissible an action that would otherwise be irrational. One reason has greater requiring or justifying strength than another when the first can overcome all the opposing reasons that the second can, and more: it has greater requiring strength when it can render an action irrational despite those opposing reasons, and greater justifying strength when it can render an action rationally permissible despite those opposing reasons. Cases like Case One are possible because reasons can justify more than they require. That the bike ride would be more enjoyable makes it more strongly justified, but it does not make it required. It is not the case that I ought to go, and not irrational to stay at home. In Case Two, by contrast, the reason not to wreck my career has requiring strength, and the fact that resigning will relieve my current stress does not have enough justifying strength to prevent the action from being irrational.

 We might wonder what these two relations – justifying and requiring – have in common. Why think of “reasons for action” as a genus of which these are two different species? But that is not hard to answer. What reasons for action have in common is that they are answers to “Why?” questions about action. The difference between justifying and requiring reasons comes from the difference between the questions they contribute to answering. The former are potential contributions to answering “Why may I do that?”; the latter, “Why must I do that?”

 So this view supplies us with a coherent explanation of the difference between Cases One and Two. It also allows for cases of supererogation in which reasons of altruism can justify but not require making significant personal sacrifices to help other people ([Gert 2004: Ch.2](#_ENREF_10)). However, in doing so it makes a big departure from the Simple Picture. It abandons the idea of *the* weight of reasons on which the Simple Picture is built.

 However, before endorsing this departure we ought to ask whether it really is necessary. Is the Simple Picture’s treatment of these cases obviously wrong? It can treat cases of supererogation as instances of incommensurability – cases in which the reason I have to help other people is neither stronger nor weaker than the reason to avoid the sacrifice of my own interests, but nor are they equally strong.[[9]](#endnote-9) And it can maintain that the difference between Cases One and Two is only a difference of degree. In Case One, there is most reason to go on the bike ride. So I ought to go, and it is irrational not to. After all, I am looking for something enjoyable to do, and the bike ride will be more enjoyable. It is irrational for me to forgo more enjoyment in favour of less. But it is only mildly irrational; wrecking my career in Case Two is much more so. And while in both cases, I ought to do what there is more reason to do, the “oughts” in the two cases are of different degrees of seriousness, thanks to the different strengths of the reasons involved.

**V: Rational Deliberation**

Both the Simple Picture and Gert’s alternative to it make the connection between rationality and the recognition of reasons a very tight one. The Simple Picture characterizes rational deliberation and action as responding correctly to the weights of our reasons. Rational deliberation, it says, is weighing reasons correctly; acting rationally is doing what you have sufficient reasons to do. Gert’s alternative explains the justifying and requiring strength of reasons by reference to what it is or is not irrational to do.

 But the connection between rationality and the recognition is not as tight as either of these views claims. Just as we can be rationally mistaken in our beliefs about anything else, we can be rationally mistaken in our beliefs about objective reasons ([Korsgaard 1986: 11-12](#_ENREF_18)). If my wife’s trick is convincing enough, it could be rational to decide to take her to hospital, and to take her there: that is because her behaviour gives me good evidence that there is a decisive objective reason to do so (namely that she has gone into labour). The fact that there is actually no such reason does not stop my deliberation and action from being rational. One can rationally believe that there are good reasons for performing an action, and rationally act on those beliefs, despite their being false.

 Conversely, one can act irrationally while doing what there is in fact a decisive reason to do, and even while assigning the correct weights to the reasons bearing on one’s action, in accordance with the evidence. Delusional deliberation or obsessional action does not become rational on the rare occasions when an accordance of those kinds happens by fluke. Rational deliberation and action requires a sensitivity to that evidence and those beliefs, and guidance by this sensitivity. So a better suggestion about the relationship between reasons and rationality is this:

Acting rationally is non-accidentally doing what you have a good justification for believing you have sufficient reasons to do. Rational deliberation is deliberation that is non-accidentally responsive to the evidence of the weights of one’s reasons.

This sensibly allows that you can act rationally without actually going so far as to form beliefs about reasons. So it can count unreflective action as rational, as long as it is non-accidentally governed by dispositions of sensitivity to the evidence of your reasons. Likewise, rational deliberation does not need to involve thinking about all of the minor reasons that have some bearing on your action; but it needs to be deliberation which, again, is non-accidentally governed by dispositions of sensitivity to the evidence of your reasons.[[10]](#endnote-10)

 Judgements of rationality are evaluative: they are a kind of judgement of goodness or badness. What such judgements evaluate, according to the Simple Picture, is the quality of our responsiveness to reasons. The proposal just offered retains this broad idea, but refines it. To appreciate the rationale for the refinement, we can ask: what are the standards for evaluating that responsiveness that there are good reasons for us to hold each other to?[[11]](#endnote-11) What standards does it make sense to aspire to, to be disappointed at not meeting, to encourage others to improve themselves towards attaining, to regard as a measure of personal excellence or deficiency? If we compare, as answers to this question, the standard of perfect conformity with the reasons themselves, and the standard of non-accidental sensitivity to the evidence of the reasons, the latter is more reasonable. The former is useless as an evaluative standard, since adopting it could not help to guide us to comply with it.

 Can our proposal be improved further? We can consider three objections. First, some philosophers maintain that the rationality of my action depends on my beliefs about the reasons, and not on the justification I have for such beliefs ([Scanlon 1998: 23](#_ENREF_29); [Kolodny 2005: 554](#_ENREF_17)). So if I form some wildly irrational belief about what I have sufficient reasons to do, and then do it, I thereby act rationally. What is irrational here is my belief, not my action: I am theoretically but not practically irrational. In support of this, it can be pointed out that, having formed my irrational belief, the failure to act on it would be a further form of irrationality. I would then be irrational both in the epistemic route to my normative belief and the practical response to it.

 However, that line of thought runs into the following problem. It treats normative belief-formation as a process that is completed prior to supplying an input to decision, with the rationality of action depending only on how one goes on to process that input. But certain kinds of post-hoc rationalization do not fit that picture. Sometimes, the irrational determination to perform an action recruits normative thought to produce convenient judgements: you can convince yourself that it is wise to buy the motorbike that the glamorous sales assistant has skilfully insinuated as the cure to your mid-life anxieties. In this kind of case, your normative beliefs do not function as independently formed inputs to decision. The converse process is occurring: your irrational decision is shaping your normative thought. This makes it misleading to describe this as a case that combines theoretical irrationality with practical rationality. Here, your practical *ir*rationality – your determination to press ahead with the action regardless of the reasons that really bear upon it – causes your theoretical irrationality. Thus, the conformity of one’s actions with one’s normative beliefs, even if non-accidental, is not enough for practical rationality, since one distinctive form of practical irrationality is the determination to engineer just such an agreement, come what may. This makes it preferable to stick with the more restrictive condition on rational action imposed in the proposal we have above.

 The other two objections to this proposal, though, do force us to revise it. The second comes from cases that combine akrasia with instrumental incompetence. Sometimes (often?) our decisions are akratic. We believe we ought not to do something and decide to do it anyway. Those decisions are irrational.[[12]](#endnote-12) But when we make them, incompetence in implementing them can amount to a further kind of irrationality. Suppose I can see that I ought not to chop down the tree in my garden, but akratically decide to do so. I might then pursue a ridiculous means – try to chop it down with a knife and fail. Now I am doubly irrational, both in my akratic decision and my instrumental incompetence. I am less rational than an equally akratic but instrumentally competent person would have been. The proposal so far fails to recognize this. It makes deciding to chop down my tree with a knife no less rational than deciding to chop it down with a chainsaw, since neither decision would be responsive to the evidence of my reasons, which tell me not to chop it down.

 To solve that problem, we need to allow for the way in which it can make sense to evaluate instrumental deliberation as more or less rational when it is not sensitive to the evidence of one’s reasons. To do so, we can say this. In a case like the one above, my incompetence has rescued me from my akrasia. I have ended up doing what I ought – namely, not chopping down my tree. But I have done so through failing to exercise an important disposition – the disposition to take evidently suitable means to my intended ends. It should matter to me that I strengthen this disposition, because of the importance that it has elsewhere: its importance in enabling me to take the suitable means to the ends I do have good reasons to pursue. This disposition is one of a family of dispositions that are important because of the role they play in thinking about and acting on our reasons. We support these important dispositions by evaluating patterns of thought as rational or irrational. Given this importance, it makes sense to apply those standards of evaluation to the relevant patterns of thought wherever we find them – in cases in which a person is not responding to her reasons as well as those in which she is. To accommodate this point, we can revise our proposal as follows:

Acting rationally is non-accidentally doing what you have a good justification for believing you have sufficient reasons to do. Deliberation is rational to the extent that its form is such that, when it is deliberation about reasons, it is non-accidentally responsive to the evidence of their weights.

One such form of deliberation is working out the instrumentally efficient means to one’s ends. Taking obviously inefficient means to one’s ends can thereby count as irrational even when the ends, and therefore the means, are ones you can see that you lack good reasons for pursuing.

**VI: Reasons Not to Deliberate**

The third objection to this proposal is more obvious. Sometimes, rationality requires you not to deliberate. When what you should do is obvious and urgent, it can be irrational to deliberate first. If someone collapses next to you at the bus stop, pausing to identify the reasons for and against helping and assigning weights to them would itself be irrational: the obvious reasons for helping straight away are obvious reasons against doing that.

 That is one way in which thinking about your reasons can prevent you from conforming with them. There are many kinds of performance – artistic, sporting, teaching, social, medical, military – of which that is true. And beyond these, there are various further ways in which it can be true that we should not deliberate. In general, we seem not to be especially good at what psychologists call “reason-based choice” ([Shafir, Simonson et al. 1993](#_ENREF_34); [Mellers, Schwartz et al. 1998](#_ENREF_21)), so we are often better off making decisions by other methods. Sometimes, when we cannot help thinking about a fact, we should disregard its weight as a reason. If you are a competitive athlete, it may make sense not to count the pain and discomfort experienced in training as reasons to stop, and instead simply obey your coach. Here, you may need to combat your tendency to exaggerate the strength of certain reasons by discounting them. Sometimes, there may be nothing wrong with thinking about a reason and its weight, but we still should not be guidedby it in acting ([Raz 1999: Ch. 1 and Postscript](#_ENREF_24)). According to Patricia Greenspan ([2005](#_ENREF_12)), some reasons permit (but do not require) us to leave others out of our deliberation about what to do.[[13]](#endnote-13) As Nomy Arpaly points out, deliberation can be a way of talking yourself out of responding to what really is important ([2003: 25-9](#_ENREF_1)). There can be circumstances in which you should not believe a proposition that is a reason; not believe that it is a reason; believe falsely that it is not a reason;[[14]](#endnote-14) and so on.

 That gives us various different ways of “not deliberating”, and at least five different kinds of reasons for not doing so. These can be reasons of (i) dominance, where one reason for action is so obviously decisive that it would be ridiculous for others to feature in thinking about what to do; (ii) non-disruption, where deliberation would distract from fluent performance; (iii) fallibility, where proneness to deliberative error makes non-deliberation a more reliable way of acting as one ought; (iv) obedience, where non-deliberation is commanded by legitimate authorities; or, more unusually, (v) self-manipulation, where large incentives attach to deliberation itself rather than the resulting action. In cases of each of these kinds, saying that one should not deliberate could mean one of two things. It could mean that there are decisive objective reasons against deliberating. Or, when the evidence of such reasons is powerful enough, it could mean that it would be irrational to deliberate.[[15]](#endnote-15) Cases of the latter sort are counterexamples to the account of rational deliberation we have arrived at so far.

 It would be possible to add further complexity to this picture. But instead, it is more helpful to look at the common pattern. The key to doing that lies in the observation: deliberative thought about one’s practical reasons is itself a kind of action on which practical reasons bear, for and against. Given this, everything we say about reasons for action and the rationality of action can be applied to deliberation as well. The fact that an action would be time-wasting, or distracting, or counterproductive, or disobedient, or would attract an externally imposed penalty, can provide us with a decisive reason against it. That can apply to actions of deliberating, just as it does to actions of other kinds.

 The upshot, I suggest, is this. When we give an account of rational deliberation, it needs two parts. Deliberation needs to be rational in its *form*, and that is what the description of deliberative rationality in Section V tries to capture. But it also needs to be rational as an activity, and that means satisfying same conditions as any other rational action. To accommodate this, we need to make one last modification to our account, as follows:

Acting rationally is non-accidentally doing what you have a good justification for believing you have sufficient reason to do. Deliberation is rational to the extent that (i) the action of deliberating is rational and (ii) it has a form such that, when it is deliberation about reasons, it is non-accidentally responsive to the evidence of their weights.[[16]](#endnote-16)

**VII: Exclusionary Reasons**

An “exclusionary reason” is Joseph Raz’s name for a reason not to be guided by a reason for or against an action. An exclusionary reason is not a reason against thinking about some practical reason, believing that it is a reason, or drawing conclusions about what it supports. It is a reason against *acting for* the reason it excludes.[[17]](#endnote-17) There do seem to be exclusionary reasons. The two most obvious kinds of example come from self-mistrust, when guiding your action by some important reason risks disaster; and obedience, when some other person or institution has the authority to instruct you how to guide your action.[[18]](#endnote-18)

 According to Raz, the existence of exclusionary reasons provides another objection to the Simple Picture: we should reject the principle that “[i]t is always the case that one ought, all things considered, to do whatever one ought to do on the balance of reasons.” ([1999: 36](#_ENREF_24)) We have already examined one way of doubting that principle: you might be dubious about the idea that practical reasons have a single dimension of strength. However, Raz accepts that idea. Instead, he maintains that “[i]t is always the case that one ought, all things considered, to act for an undefeated reason” ([1999: 40](#_ENREF_24)), and he makes a distinction between two ways in which a reason for action can be defeated. One way is by being outweighed by a stronger reason against the action; another is to be excluded by an undefeated exclusionary reason. To be an exclusionary reason is not itself to be a direct first-order reason for or against action, so exclusionary reasons leave the balance of first-order reasons unaffected. They do not outweigh the reasons they exclude; but they defeat them, and thereby affect what you ought to do all things considered.

 To illustrate this, we can consider one of his obedience-examples. Jeremy, a soldier, is ordered by his commander to appropriate a tradesman’s van ([1999: 38](#_ENREF_24)). Taking the van is unnecessarily intrusive: the inconvenience to the tradesman far outweighs the negligible military usefulness of his van. Even after the order is given, Raz thinks, the balance of first-order reasons may count against taking the van. However, the reasons against taking it are excluded by the authority that Jeremy’s commander exercises in giving his order. So Jeremy ought, all things considered, to take it.

 What is the point of saying that? Raz argues that this allows us to explain how, in a case like Jeremy’s, conflicting assessments apply. The order is faulty, but he ought to obey it. It is faulty because the balance of reasons counts against appropriating the van; Jeremy ought to obey it because his commander’s exercise of legitimate authority excludes the reason against doing so. This is why Jeremy can regard it as fortunate if he does his best to comply with the order but is somehow prevented from doing so. Then he has happily succeeded in both doing his best to be obedient, as he ought, and doing what he had most reason to do ([Raz 1999: 41-5, 185-6](#_ENREF_24)).

 However, this is problematic. What is the force of saying that Jeremy has done “what he had most reason to do”, if he ought not to have done it? It is hard to see how this helps to explain why Jeremy should be happy at the way things have turned out, if it carries the implication that he has failed to do what he ought, all things considered, to have done.

 One could modify Raz’s view in a way that avoids this odd implication. A distinction he makes is between conformity with one’s reasons – doing what they support – and compliance with them – being motivated to do so by your recognition of those reasons ([1999: Postscript](#_ENREF_24)). An exclusionary reason is a reason not to comply with another reason; but it is not a reason not to conform with it. Building on this, one might then make a distinction Raz does not make: a distinction between what a person ought to do and the reasons by which she ought to be guided. One might then agree with the Simple Picture that a person ought *to do* what there is most reason for her to do; but add that she ought only to be *guided* in acting by (to act *for*) non-excluded reasons. By saying that, we could retain the special role of exclusionary reasons, while making a different set of claims about what ought to be done in cases like Jeremy’s. Jeremy ought *not* to appropriate the van, but as an obedient soldier, he also ought not to act for the reasons against appropriating the van. Saying this makes more sense of how he could think that things had turned out fortunately if he obediently tried to appropriate the van but was prevented from doing so – that way, he does everything he ought.

 This variation on Raz’s view would now no longer contradict the Simple Picture, but it would complicate it. It distinguishes two different objects of all-things-considered oughts – actions, and guidance in acting – while retaining the Simple Picture’s principle that one ought to *F* just in case there is most reason to *F*. However, there remains an odd-seeming implication. Suppose Jeremy himself accepts this view, and accordingly concludes that he ought not to take the van, but ought not to act on the reasons against taking it. Where does that leave him? He could try to take it, hoping that some external intervention will prevent him, so that he ends up doing everything he ought. But that might be very unlikely. Short of that, whichever way he turns, he fails to do what he ought. How, then, can deliberation help him to decide what to do?

 One response to that would be that Jeremy’s practical situation just is normatively awkward in that way: life is not always simple. But before resorting to that, there is a more straightforward treatment to consider. Jeremy’s commander ought not to have given the order, since the balance of reasons is against it. But once the order has been given, obedience to his commander is a first-order reason for Jeremy to do as he is ordered, and it could be strong enough to be decisive. If so, Jeremy ought to obey the order and take the van. If Jeremy tries to comply with the order and is prevented, he should not feel happy about his non-compliance: he has not done what he ought. But he should still feel happy about the outcome: the tradesman’s not having his van taken avoids unnecessary inconvenience to him, and that is good. In this way, we can explain why mixed feelings make sense in a case like this without treating exclusionary reasons as a special kind of non-outweighing defeater.

 Thus, although there surely are exclusionary reasons, their existence does not seem to provide a strong objection to the Simple Picture. In either this straightforward way, or with the supplementation we considered above, the Simple Picture can provide a more satisfactory treatment of what we ought to do in such cases than Raz’s alternative.

**VIII: Undermining**

We have just encountered the idea that outweighing is not the only way in which a reason for action can be defeated. If there are others, then deliberation about our reasons for action will need to involve more than just the comparison and addition of weights, so this will be another respect in which the Simple Picture is too simple.

 One version of this idea is Raz’s claim, discussed (and resisted) above, that a reason can be defeated without altering its contribution to the balance of reasons. A different possibility worth considering is that one reason can remove another from that balance, depriving it of its weight as a reason. We can describe that as a relationship of undermining, rather than outweighing.

 To illustrate, consider this variant of Raz’s example. Suppose the van belongs to a tradesman in the same line of business as Jeremy’s brother. If Jeremy takes it, his brother’s business will get a boost. Is that a further reason for obeying the order? A view worth considering is that this fact does not add to the balance of reasons for his action at all. The fact that his brother will benefit is a reason for other actions, but not this one. Here, this fact is undermined, not outweighed: it loses its status as a reason. What does the undermining? One suggestion is this. In carrying out his duties as a soldier Jeremy acts not in his own personal capacity, but as an agent of an institution – the army – that he is committed to serving. He is performing *its* actions, not his own personal ones. So the only reasons bearing on these actions are reasons for the army to do things: the business interests of his brother do not qualify. In this way, Jeremy’s military role functions as an underminer, not an outweigher. His military role is not a reason *against* taking the van (he has been ordered to take it, after all); rather, it is a reason why his brother’s interests do not count in favour.

 That describes a position to be argued for, rather than an argument for it. To defend that reading of the example would be a much larger task.[[19]](#endnote-19) We would need to consider a broader range of examples, and assess the plausibility of the distinctions this interpretation requires us to draw. For instance, it suggests that there is a significant difference between the case just described, in which Jeremy would be helping his brother *by* exercising his military role, and one in which he can help him *instead of* exercising his military role. If, while on leave, he hears of a valuable business opportunity for his brother, ought he to delay returning from leave in order to tell his brother about it? Now he faces a choice between performing a personal action and acting as an institutional agent; so on the view just sketched, he does here need to weigh up the reasons bearing on those two different kinds of action and assess which is stronger. Beyond this, we would need to consider a wider range of other institutional actions, and evaluate the implications that this view of them would have more generally.

 This is not the place to do that. Notice that the difficulty of the task is increased by the existence of the kinds of case described in Section VI, in which a consideration should be excluded from deliberation even though it retains its weight as a reason. If there can be cases like that, we cannot establish that a consideration is undermined as a reason by showing that it does not properly feature in an agent’s deliberation.[[20]](#endnote-20) The only fully convincing way of vindicating the existence of undermining relations between reasons would be to present a general view about which such relations there are and how they are instantiated, and show that that view delivers cogent explanations of a wide range of examples.

 What general features would a view of that kind have? It would reject the sort of atomism about reasons we ascribed to Ross in Section II. If one fact can be deprived of the status of a reason by another, then its status as a reason can depend on the rest of the circumstances. However, notice that it would not require abandoning Ross’s combinatorial approach to normative theorizing – the approach that seeks to explain normative complexity through the interaction of a limited number of fundamental elements. A thought entertained by some philosophers is that it may be possible to construct a normative theory whose foundations – the sources to which we can trace the derivation of derived reasons – amount to “default reasons”, that is, facts that have the status of reasons for action unless subject to undermining. Beginnings have been made to demonstrate the explanatory coherence of such an approach, but to date little work has been done to elaborate and defend it in substantive detail ([Horty 2007](#_ENREF_14); [Horty 2012](#_ENREF_15)).

**IX: Conclusion**

I have concentrated on examining the adequacy of the Simple Picture of weighing reasons. Consequently, there are important questions about the weighing of reasons which have not been discussed here, since they are questions on which the Simple Picture is neutral. There is the metaphysical question, “What *is* it for one reason to be stronger than another?”; the conceptual question, “Is the concept of a reason more basic than the concept of ought, or vice versa (or neither)?”; and the epistemological question, “What warrants me in judging that one reason is stronger than another?” The last of these, especially, merits much more discussion than it typically receives.

 Four main revisions to the Simple Picture have been discussed. First, its talk of “summing” the weights of reasons needs to be moderated. It cannot be applied to reasons that have a common derivational source. Provided that condition is met, there is a sense in which reasons can be added: the overall support for an action is made stronger by the introduction of extra reasons. But the way in which reasons combine to determine what one ought, overall, to do cannot in general be thought of as the addition of vectors with unique cardinal magnitudes.

 Secondly, we examined whether the Simple Picture needs to be revised to allow that the action which is most strongly supported by our reasons need not be required by them. One proposal is to discard talk of *the* weight of a reason, recognizing instead two different kinds of reason-giving force: justifying and requiring. But it does not seem obviously wrong to accept the Simple Picture’s claim that we are required to do what is most strongly supported by the net balance of reasons, while allowing that those requirements have varying degrees of seriousness.

 In a third respect, however – its account of rational deliberation and action – we found that the Simple Picture is definitely too crude. It says that rational deliberation is weighing reasons correctly, and that acting rationally is doing what there is most reason to do. Instead, we should characterize rationality as non-accidental sensitivity to evidence of one’s reasons.

 Fourthly, we considered the possibility that the normative weight that one reason-giving fact contributes to supporting an action may be affected by its relationship to the other facts that are part of the same context. This does not cast doubt on the idea that the overall normative character of our actions is determined by the comparative weights of the reasons that bear on them, nor that rational deliberation involves doing our best to assess those weights. But it suggests that good deliberation may also first need to be sensitive to the ways in which the interactions between the different elements of a situation may determine which of those elements are reason-providing, and how much weight they carry.[[21]](#endnote-21)

Word count

text 8,981

notes and refs 1,060

total 10,041

**References**

Arpaly, Nomy (2003). Unprincipled Virtue. New York: Oxford University Press.

Broome, John (1997). "Is Incommensurability Vagueness?". Incommensurability, Incomparability and Practical Reason. R. Chang. Cambridge, Mass.: Harvard University Press**:** 67-89.

Broome, John (2004). Weighing Lives. Oxford: Oxford University Press.

Broome, John (2013). Rationality through Reasoning. Oxford: Wiley Blackwell.

Chang, Ruth (2002). "The Possibility of Parity". Ethics **112**: 659-688

Cullity, Garrett (2008). "Decisions, Reasons and Rationality". Ethics **119**: 57-95

Cullity, Garrett (2013). "The Context-Undermining of Practical Reasons". Ethics **124**: 1-27

Dancy, Jonathan (2004). Ethics without Principles. Oxford: Clarendon Press.

Dancy, Jonathan (2006). "Enticing Reasons". Reason and Value: Themes from the Moral Philosophy of Joseph Raz. R. J. Wallace, P. Pettit and S. Scheffler. Oxford: Oxford University Press**:** 91-118.

Gert, Joshua (2004). Brute Rationality. Cambridge: Cambridge University Press.

Gert, Joshua (2007). "Normative Strength and the Balance of Reasons". Philosophical Review **116**: 533-562

Greenspan, Patricia (2005). "Asymmetrical Practical Reasons". Experience and Analysis: Proceedings of the 27th International Wittgenstein Symposium. M. E. a. J. C. M. Reicher. Vienna: oebv&hpt**:** 115-122.

Griffin, James (1986). Well-Being: Its Meaning, Measurement and Moral Importance. Oxford: Clarendon Press.

Horty, John F. (2007). "Reasons as Defaults". Philosophers’ Imprint **7**. <[www.philosophersimprint.org/007003/](http://www.philosophersimprint.org/007003/)>

Horty, John F. (2012). Reasons as Defaults. New York: Oxford University Press.

Jackson, Frank (1981). "Decision-theoretic Consequentialism and the Nearest and Dearest Objection". Ethics **101**: 461-482

Kolodny, Niko (2005). "Why be Rational?". Mind **114**: 509-563

Korsgaard, Christine M. (1986). "Skepticism about Practical Reason". The Journal of Philosophy **83**: 5-25

Lewis, David (1986). "Causal Explanation". Philosophical Papers. his. Oxford: Oxford University Press**:** 214-240.

McNaughton, David and Piers Rawling (2004). "Duty, Rationality, and Practical Reasons". The Oxford Handbook of Rationality. A. R. Mele and P. Rawling. Oxford: Oxford University Press**:** 110-131.

Mellers, B. A., A. Schwartz, et al. (1998). "Judgment and Decision Making". Annual Review of Psychology **49**: 447-477

Moore, G.E. (1903). Principia Ethica. Cambridge: Cambridge University Press.

Parfit, Derek (2011). On What Matters. Oxford: Oxford University Press.

Raz, Joseph (1999). Practical Reason and Norms. Oxford: Oxford University Press.

Raz, Joseph (2000). "The Truth in Particularism". Moral Particularism. B. Hooker and M. O. Little. Oxford: Clarendon Press**:** 48-78.

Raz, Joseph (2006). "The Trouble with Particularism (Dancy’s Version)". Mind **115**: 99-120

Regan, Donald (1997). "Value, Comparability, and Choice". Incommensurability, Incomparability and Practical Reason. R. Chang. Cambridge, Mass.: Harvard University Press**:** 129-150.

Ross, W.D. (1930). The Right and the Good. Oxford: Clarendon Press.

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

Schelling, Thomas C. (1960). The Strategy of Conflict. Cambridge, Mass.: Harvard University Press.

Schroeder, Mark (2011). "Holism, Weight, and Undercutting". Nous **45**: 328-344

Schueler, G.F. (1979). "Exclusionary Reasons". The Personalist **60**: 407-410

Setiya, Kieran (2007). Reasons without Rationalism. Princeton: Princeton University Press.

Shafir, Eldar, Itamar Simonson, et al. (1993). "Reason-Based Choice". Cognition **49**: 11-36

Stratton-Lake, Philip (2011). "Eliminativism about Derivative Prima Facie Duties". Underivative Duty: British Moral Philosophers from Sidgwick to Ewing. T. Hurka. Oxford: Oxford University Press**:** 146-165.

**Notes**

1. So, more strictly, the Simple Picture sees the reason-relation as a scalar, not a vector. It only has a magnitude, which can be either positive or negative. [↑](#endnote-ref-1)
2. This is what we are primarily interested in. But when we do not have enough information to figure that out, what we are secondarily interested in is what is the rational thing to do given the evidence. And this can sometimes be an action which we know is not supported by the strongest objective reason. (Frank Jackson ([1981: 463](#_ENREF_16)) illustrates this with the example of a doctor who knows that drug A will partly cure her patient, and that one of drugs B and C will completely cure him and the other will kill him, but does not know which.) [↑](#endnote-ref-2)
3. And if this, in turn, derives its normative significance from some further, more fundamental consideration, one might want to say that only the latter consideration is the real reason. Compare the discussion of “complete reasons” in ([Raz 1999: 22-5](#_ENREF_24)) and ([Raz 2000: 58-61](#_ENREF_25)); also ([Stratton-Lake 2011](#_ENREF_35)). [↑](#endnote-ref-3)
4. Paul Weirich helped me to appreciate this. [↑](#endnote-ref-4)
5. See ([Dancy 2004: Ch. 5](#_ENREF_8)) – for his rejection of Moorean non-compositionality, see Ch.10. [↑](#endnote-ref-5)
6. See ([Ross 1930: Ch. 2](#_ENREF_28)); for discussion, see Philip Stratton-Lake’s introduction to the 2002 edition, pp.xxxiii-xxxviii. [↑](#endnote-ref-6)
7. There are four other alternatives to consider, though. One is the epistemic view that the difficulty of comparing meals with concerts means that although the reason to go to Angela’s must be stronger than, weaker than or equal to the reason to go to the concert, we cannot know which ([Regan 1997](#_ENREF_27)). The second is the semantic view that the comparisons are vague, not incommensurate – giving us instead the conclusion that we cannot say that either is stronger nor that they are equal ([Broome 1997](#_ENREF_2)). The third is that the meal and concert are “on a par” – treating that as a comparability-relation that is distinct from the relations of stronger than, weaker than and equal to ([Chang 2002](#_ENREF_5)). And a fourth alternative is to avoid incommensurability by embracing holism. Given only the choice between Angela’s and the Chopin concert, the two reasons are equal in strength; given only the choice between Bernie’s and the concert, the reasons are equal in strength; but if you have a choice between all three, the strengths of the reasons change – now the reasons to go to Angela’s and the concert are equal in strength, but the reason to go to Bernie’s is weaker. [↑](#endnote-ref-7)
8. However, perhaps we can still assign to each reason a range of numerical values ([compare Broome 2004: Ch. 12](#_ENREF_3)). [↑](#endnote-ref-8)
9. Gert claims that his view allows us to “explain the phenomena that seem to support incommensurability without leading to the obscurities that surround that notion.” ([2007: 549](#_ENREF_11)) However, this does not seem to be generally true. Sometimes, it seems that one’s reasons can require choosing A+ over A and B+ over B, but permit choosing B over A+ or A over B+. (E.g. A is a state in which my pain is lessened and in A+ it is lessened further; B is a state in which my lucidity is improved and in B+ it is improved further.) There is no consistent assignment of requiring and justifying strengths to the reasons favouring these options that can deliver this result (unless one resorts to the fourth, holistic alternative mentioned in note [7], making the assignment of strengths to reasons circumstance-relative). [↑](#endnote-ref-9)
10. It is a good, further question what exactly has to be true of me for my responsiveness to the evidence of my reasons to count as “non-accidental”. [↑](#endnote-ref-10)
11. This way of thinking about the evaluative standards that govern rationality-judgements – I call it a “standard-fixing account” of the relation between reasons and rationality – is developed further in ([Cullity 2008](#_ENREF_6)). [↑](#endnote-ref-11)
12. One problem with the proposal currently under discussion is that it sometimes fails to deliver that result. Huckleberry Finn is apparently sensitive to the evidence of his reasons in what he decides to do; but he forms the wrong beliefs about those reasons and then akratically fails to act on his beliefs. That does seem to be a kind of irrationality. The revised proposal at the end of this section deals with this sort of case: see note [15]. [↑](#endnote-ref-12)
13. Like Gert, Greenspan explains a reason’s requiring an action in terms of the appropriateness of criticism of the agent as irrational ([Greenspan 2005: 387-9](#_ENREF_12)). [↑](#endnote-ref-13)
14. This might happen in a “Schelling case”, where the most effective way to prevent someone from threatening you is to convince him you are irrational. See ([Schelling 1960](#_ENREF_30)). [↑](#endnote-ref-14)
15. Then it is a subjective “should” of the kind we met in Section I. [↑](#endnote-ref-15)
16. Notice how this now deals with cases of “inverse akrasia” such as Huckleberry Finn ([Arpaly 2003: 8-11](#_ENREF_1)). Huck wrongly believes he ought to turn Jim over to the slave-hunters, but he fails to do so out of responsiveness to the reasons against doing so. So his action counts as rational, on the account we have arrived at, because of its non-accidental responsiveness to the evident reasons. But his deliberation is irrational, because of its akratic form. [↑](#endnote-ref-16)
17. As G.F. Schueler ([1979](#_ENREF_32)) points out, there is a sense in which every reason against an action is a reason against acting for the reasons in favour of that action. But Raz’s idea is that for a consideration to function as an exclusionary reason is not for it to function as a contrary first-order reason. It is to be a reason against acting *for* a first-order reason – against being motivated or guided by that first-order in acting ([Raz 1999: 178-85](#_ENREF_24)). [↑](#endnote-ref-17)
18. The first kind of case is illustrated by Raz’s example of Ann, who is given the opportunity to make an investment decision with a deadline of midnight, but judges herself to be too tired and upset to make an accurate decision. The second is illustrated by the examples of Colin, who has promised his wife to consider only his son’s interests in making decisions about his schooling, and the soldier Jeremy, discussed in what follows ([1999: 37-9](#_ENREF_24)). [↑](#endnote-ref-18)
19. I attempt part of that task in ([Cullity 2013](#_ENREF_7)). [↑](#endnote-ref-19)
20. Here, I am disagreeing with Mark Schroeder’s claim that “facts about the weight of reasons are facts about correct deliberation” ([2011: 336](#_ENREF_31)). [↑](#endnote-ref-20)
21. For help in improving earlier drafts, I am grateful to Antony Eagle, Jordi Fernandez, Philip Gerrans, Joshua Gert, Frank Jackson, Jon Opie, Philip Pettit, Daniel Star and Paul Weirich. [↑](#endnote-ref-21)