

Minimalism about Intention: A Modest Defence

Sergio Tenenbaum

Introduction

Intentions have been recently given the starring role in theories of practical rationality. Michael Bratman, for instance, has argued for various requirements governing intentions, especially future-directed intentions.¹ At the same time, there has been a general tendency to formulate more traditional principles of practical rationality in terms of intentions. Intentions, but not actions, seem to be under the rational control of the agent.²

My general view is that intentions should be playing a far more modest role in the theory of practical rationality; the main role should be given back to actions. However my goal here is just to take one step in arguing for this view. I first explain and motivate the general deflationist view of intention I am sympathetic to; I call it ‘minimalism’ about intentions. Minimalism contrasts with more robust views of intention, and in particular views that imply the existence of intention specific rational requirements; that is, rational requirements governing intentions that are not simple consequences of rational requirements on actions. I then distinguish four different types of putative rational requirements that are incompatible with minimalism. I argue that three of these requirements are supposed to be based on what I call ‘internal’ grounds.³ I then argue that these grounds are incapable of justifying any kind of rational requirement that is incompatible with minimalism. Detailed examination of the fourth type of putative requirement is left for another occasion, as it is supposed to be based on an entirely different kind of justification (what I call ‘external’ grounds). But I do end by briefly sketching some reasons to be skeptical that this kind of justification would succeed. In sum, this paper tries to show that minimalism about intention turns out to be a rather plausible and compelling view.

¹For a recent example, see Bratman (2012).

²See, for instance, Broome (2013) for a formulation of the classical instrumental principle in terms of intentions.

³I am grateful to Sarah Paul for suggesting, in a different context, this label (and a more natural way of dividing the options here)

Minimalism about Intentions

When I act, there is at least one intention with which I act, and the rational requirements on my actions are sensitive to what I intend in acting, or to the end of my action. So, if I am fixing the car (intentionally) by cleaning the oil pump, or, in other words, if I am cleaning the oil pump with the intention of fixing the car, then whether I am being instrumentally rational will depend in part on whether my actions are (or at least on whether I reasonably expect them to be) conducive to my end of fixing the car. Secondly, if I carry out my action of fixing the car to its conclusion and nothing untoward happens, I have fixed my car intentionally. But even if I don't, if my skills prove unequal to the task or if my vehicle is beyond repair, it still true that I intended to fix the car. More importantly, this failure of my intention is compatible with my being perfectly rational. Some philosophers might want to deny some of these claims, but I will take them to be desiderata for any theory of rationality. Minimalism, as I understand it, does not deny these platitudes. Instead, minimalism about intentions denies that *future-directed intentions*(FDIs), intentions that are not guiding one's present actions, give rise to any non-derivative practical requirements.⁴ More particularly, minimalism accepts the following:

1. The Primacy Of Acting With an Intention

Bratman has famously argued that the old focus on intention in action obscured the "second face" of intention; namely, future-directed intentions. FDIs are supposed to give rise to their own requirements. Minimalism denies exactly this point; according to a minimalist view, only intentions in action generate rational requirements. To the extent that adopting an FDI generates a requirement, it must be an instance of a requirement on intention in action or some other requirement of practical rationality (see below). In fact, we need not mention an intention at all in formulating such requirements; we can simply talk about requirements that apply to *pursuing an end*.

2. The Primacy of Evaluative or Normative Judgments

Minimalism accepts that one can be under various requirements that stem from various normative and evaluative judgments. According to

⁴As I hope will soon be clear, this does not mean that they play no role in the life of the rational agent. The claim is only that they do not generate any special requirements on action. Similarly, supposition plays a very important role in forming beliefs, but, arguably, supposing that **p** generates no special requirements on belief formation.

minimalism, requirements on FDIs must derive from requirements on normative or evaluative judgments that potentially justify such FDIs.

Why would anyone be attracted to a minimalist position? First, there is the obvious issue of theoretical simplicity. As long as we accept that there are genuine rational requirements, it is hard to deny that some of these govern evaluative or normative judgments or beliefs and some of these govern actions or at least intentions in action. Since we are already stuck with these two types of requirements, it would be better if we could avoid others.

More importantly, the thought that the paradigmatic case of practical inconsistency is incoherence in action seems extremely plausible. Let us take the following passage from Anscombe:

the contradiction of ‘I’m replenishing the house water-supply’ is not ‘You aren’t, since there is a hole in the pipe’, but ‘Oh, no, you aren’t’ said by someone who thereupon sets out e.g. to make a hole in the pipe with a pick-axe.(Anscombe 2000, p. 55)

Whatever Anscombe is saying here, it seems right that the paradigmatic case of practical incoherence would be for me to be engaged in these two actions at the same time, adjusting the pipe with one hand while making holes in it with the pick-axe with another.⁵ If intentions do not result, or even threaten to result, in such actions, and do not generate, or at least threaten to generate, a conflict between our actions and our judgments about how we should act (or what would be good to do), then it is not clear why conflict of intentions would amount to *practical* incoherence.⁶

⁵Doubtless, one could be coherently engaged in actions that would result in similar movements. The incoherence is being engaged in *replenishing the water supply* while at the same time making holes in the pipe (and while knowing the basic hydraulic facts).

⁶Recently, some philosophers have defended the view sometimes dubbed “cognitivism” about instrumental rationality (See Wallace (2001) and Setiya (2007)). Cognitivism in this sense claims that instrumental requirements are at bottom theoretical requirements; if I intend to φ , but I believe that I cannot φ unless I ψ , and I do not intend to ψ , then if intending to φ requires the belief that I will φ , it seems that I hold inconsistent beliefs; namely that I will φ , that I cannot φ unless I ψ , and that I will not ψ . But this literature often focuses on formulating the requirement in terms of intention; even if intentional action requires belief it is not clear that this kind of reduction holds much promise for the case of action. For suppose that I am now trying to fix my car, but I think that the prospects of success are dim. We can now grant, for the sake of the argument, that if I do not believe that I can fix it, I cannot be said to be intending to fix my car; perhaps I can only say that I am trying to fix the car. Still, having the car fixed is the end of my action, and it would be practically incoherent if I were at the same time pouring acid into the engine, which would make fixing the car nearly (though not metaphysically) impossible.

Finally, despite appearances, whether the intention results in action or not *is* in fact relevant for the rational evaluation of the agent. It seems that, if I intend to go to Paris, but weather patterns cancel my flight, I am just as rational as if everything went well and I got to Paris. But this is not literally true. Whether I am ‘just as rational’ depends on what I do instead of going to Paris once I find out about the flight cancellation. If I proceed to the gate as if nothing had happened, or if I take the plane to Madrid instead, I do not count as rational. In fact, in most such cases, rationality would require that I make other arrangements to get to Paris (like, possibly, taking the next plane) and proceed with my action of going to Paris. Moreover, going to Paris would require many more actions from me, such as taking a seat on the airplane, leaving the airplane, taking a cab from the airport to the desired location in Paris, etc. The presence of the intention up to the time of the flight cancellation neither guarantees that I am acting rationally now, nor that what I have done in terms of the exercise of my rational powers is the same as in the case in which I go to Paris. This is an instance of a much more general phenomenon; whatever I intend, if nothing interferes with my intention, I carry out in action. Since I cannot foresee every detail of the circumstances in which I act, whether I act rationally or not will depend not only on my previous intention but also on how such an intention is carried out in action.⁷ Of course, with a bit of philosophical ingenuity, we can come up with all sorts of cases of paralysis and illusions, in which many things are intended but no action is undertaken by no fault of the agent. But it does not follow from this observation that practical requirements govern intentions rather than actions.

These considerations are only intended as preparatory remarks. Obviously, there is much that can be said in response to these claims, and there are reasons to favour intentions in formulating principles of practical rationality that I have not formulated here⁸; I only aim to show that minimalism might be a cause worth fighting for.

However, it is not clear how we could attribute to me any theoretical incoherence if I were to proceed in this fashion. Some versions of the “belief requirement” that replace beliefs with credences (see, for instance, (Setiya 2008)) might avoid this difficulty, but they are problematic in other ways (ref to Knowing the Good and ...)

⁷For a more detailed argument for this point, see my [ref to Conclusion of PR

⁸Most notably the fact that intentions, but not actions, seem to be within direct control of the agent. I think this view is incorrect, but I cannot argue against it here.

Squeezing Intentions Out

Given what we said above, the obvious strategies to “squeeze” intentions out of a theory of practical rationality would be to show that certain requirements that seem to be requirements generated by FDIs are either requirements generated by certain evaluative judgments or by certain actions that one is currently undertaking. Elsewhere, I have argued that policies and projects are best understood as ordinary extended actions, at least for the purposes of a theory of practical rationality;⁹ and that requirements governing policies, projects, etc. are really ordinary, instrumental requirements that apply to any action that extends through time.¹⁰ Also, it is quite clear how, in some cases, adding special intention requirements would be superfluous, given that the normative judgments that underwrite the relevant intention already generate the same requirements. For instance, if I judge that I should be in Paris in November,¹¹ I would, insofar as I am rational, form the intention to be in Paris in November. But if I can’t be in Paris in November unless I buy tickets in October, I am required to buy the tickets in October. Here to say that such a requirement is a consequence of forming the intention to go to Paris in November is superfluous. After all, you were already under this requirement given that you judged that you ought to be in Paris in November.¹²

Another option open to the minimalist is to try to show that requirements generated by FDIs that cannot be accounted for by actions or judgment requirements are merely apparent. It is worth looking at a concrete example to get a better sense of when this “explaining away” strategy is most likely to be used. Let us suppose I judge that it would be good to go on vacation

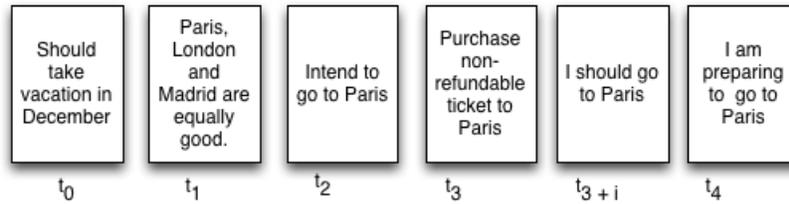
⁹Tenenbaum (in preparation).

¹⁰Arguably, these are all the actions there are, but, at the very least, the vast majority of our actions are extended actions.

¹¹I am assuming that all agents’ judgments discussed are correct (unless otherwise indicated) to avoid complications relating to how such requirements apply in cases of error.

¹²Complications arise if certain kinds of irrational behaviour creep in. For instance, suppose I judge that I ought be in Paris in November, but I know that, given my fear of flying, I would never get into a plane, and thus (assuming that other means of transportation are unavailable) never get to Paris. In this case, I am (arguably) under no requirement to buy tickets to Paris in October. One might take this to be evidence that the requirement is generated by the FDI not by the judgment. But this can’t be right. After all, if ignoring my fear of flight I went ahead and formed the intention to go to Paris in November, I would still be, on the same grounds, under no requirement to buy the useless tickets. The more plausible view is that my fear of flight is a reason to revise my judgment that I ought to go to Paris in November.

in December. There are various destinations that interest me, so I start deliberating about where I should go. I conclude that various destinations would be equally good, but, of course, I must settle on one of them. I finally form the intention to go to Paris. I wait a few days till I get my paycheck and I buy a nonrefundable ticket to Paris. As December approaches, I do various things in preparation to my going to Paris. So here is the temporal progression of my judgments, actions, and intentions:



Since no intention is formed prior t_2 , all the relevant requirements (if any) must be generated by the normative judgment I accepted in t_0 . But it will also be hard to create problems for the minimalist by focusing on what happens after t_{3+i} , since the minimalist can appeal to the normative judgments formed (and the actions undertaken) from this point on to explain any putative requirements. Only requirements that supposedly “kick in” at some point between t_2 and t_3 would seem to pose problems for the minimalist. What exactly would those requirements be? Here are a few possibilities:

[NONREC] An agent should be disposed not to reconsider her intention, unless certain special circumstances obtain.

[PERSIST] An agent should not abandon her intention without reconsidering unless certain special circumstances obtain.¹³

[NOGRAT] Reconsideration alone should not suffice to abandon one’s intention unless certain special circumstances obtain (No gratuitous change of mind).

[MEANS] If you intend the end, you must intend the means unless certain special circumstances obtain.

The clauses “unless certain special circumstances obtain” are left here for our opponent’s benefit; I will not try to spell them out further. There are various motivations to adopt each of these requirements, but I would like to distinguish two broad categories of grounds to accept them.

¹³Broome (2013) defends a more precise version of this principle. See also Bratman (2012) for a similar principle.

We can think that certain grounds for accepting a rational requirement flow directly from the nature of rational agency or the state in question; they are, as I will call them, *internal* to rational agency as such. For instance, [MEANS] is clearly a case in this category; Kant goes so far as to say that the similar hypothetical imperative is *analytic*, that a perfectly rational agent wills the necessary means whenever she wills the ends. Whether or not this principle is analytic, it does seem to be a principle whose justification would not appeal to anything beyond the nature of an intention and the nature of rational agency. [NONREC] on the other hand is probably best supported by *external* considerations; that is, considerations that appeal to the limits and obstacles encountered by the particular circumstances in which a particular set of rational agents (human beings) exercise their rational capacities. For instance, in defending a special role for FDIs in a theory of rationality, Holton says:

What need do we have of states that are stable and controlling in this way? ... [W]e are epistemically limited creatures. Information is scarce, and costly to obtain. Reasoning on the information that we have takes time and effort. It is rational then to allocate our scarce resources by limiting the amount of time we spend looking for information and reasoning on it. ... [And] we need some way of storing our decision so that we act on it tomorrow without reconsidering it (for to reconsider it would violate the requirement that we not reason about it further). We need an intention. (Holton 2009, p. 3.)

Here, I will be concerned mostly with internal grounds for accepting such requirements.¹⁴ Both [PERSIST] and [NOGRAT] can be given internal justifications; one might argue that it is of the nature of forming intentions for reasons (or in a rational manner) that one cannot give them up without adequate reason. Given that they are amenable to very similar kinds of justification, I'll often refer to [PERSIST] and [NOGRAT] together as "intention persistence requirements".¹⁵ In the end, it will turn out that there is no good justification for [PERSIST] and [NOGRAT], internal or external. I will also argue that if one cannot make a case for those requirements, one cannot make a case that there is an independent instrumental requirement like [MEANS] that applies non-derivatively to FDIs either. If the argument succeeds, only [NONREC] is left standing.

¹⁴Elsewhere, I argue against the main external grounds for accepting special requirements on intention. See (Tenenbaum in preparation).

¹⁵For a recent defence of such requirements, see ?.

Persistence in Belief

It is worth asking whether similar persistence requirements may apply to beliefs. Doubtless, some beliefs tend to persist in our mental economy. It is hard to imagine how one could gain progressively more knowledge if beliefs were to evaporate as soon as they were formed. I would boldly hazard an armchair guess that a significant subset of our beliefs persist, in normal cases, for a significant extent of time. But can this important role that belief preservation plays in our lives give rise to normative requirements governing belief persistence? Let us start with an obviously false normative constraint:

[OFC] It ought be the case that if S believes that \mathbf{p} at t_0 then S believes that \mathbf{p} at t_1 .

There are two obvious reasons why [OFC] is false. First, we can gain new information. But we can also lose information by simply forgetting certain things.¹⁶ Although one might be tempted to think that forgetting is a case in which we fall short of full rationality,¹⁷ this is implausible. We often forget old phone numbers, security codes we no longer need, etc. This kind of forgetfulness seems to manifest a rational disposition; it would certainly be a shortcoming for a rational being to overload her cognitive capacities with useless information. A more reasonable requirement of persistence would take these cases into account; it would be an on-off belief version of some kind of prohibition on non-trivial updating on no new information. Here is a possible version of such a principle:

[NOINVERSION] A rational believer does not ‘update’ from believing \mathbf{p} to believing **not** \mathbf{p} , unless she has new information, etc.

I think some improved version of this principle is correct. But it is not a special requirement on persistence of belief; rather it is a consequence of how the basing of a belief persists. It is easier to see this point, if we assume that, typically, if a belief is preserved so is its basing. If I have accessible to me evidence that is sufficient as a basing for \mathbf{p} , I should typically not form the belief that **not** \mathbf{p} . But even if basing is not preserved when belief is preserved, as long as we still want to say that memory preserves justification,

¹⁶In fact, since this is an impersonal principle, it would seem to be massively violated by the subject’s death. I’ll leave these complications aside.

¹⁷In an early version of ?, Broome argues that forgetting something one previously believed is a failure of rationality. My understanding is that Broome no longer holds this view.

we must accept that my memory somehow provides justification for **p**. My memory will in this case do the same work of preventing me from rationally forming a belief that **not p** that the preservation of basing would have done. In either case, the availability of basing for **p** will, in normal circumstances, prevent me from rationally forming a belief that **not p**.

A stronger principle might have a better shot of being a nonderivative principle of belief persistence:

[NOSWITCH] A rational believer does not ‘update’ from belief that **p** to suspending belief about **p**, unless she has new information, forgets, etc.

Something like [NOSWITCH] seems more likely to be a nonderivative principle. However, [NOSWITCH] is either superfluous or false. Let us suppose that there is no permissible ‘band’ such that, given the same evidence, it would be equally rational either to suspend judgment about **p** or to believe that **p**. In this case, [NOSWITCH] turns out to be true, but it is a simple consequence of the fact that my evidence determines a single rational epistemic attitude. In other words, unless I have new information, forget, etc., the evidence at my disposal is the same, and thus the unique rational attitude that I should adopt is also the same.

Things are a bit more complex if we allow that in some case it is possible that the evidence does not fully determine one’s rational attitude towards a proposition. Perhaps in some cases, it is equally rational to belief or suspend judgment with relation to **p**.¹⁸ If this were the case, one might be tempted to argue that [NOSWITCH] is a nonderivative valid requirement. Suppose a jury is divided in a difficult case. It seems that the mere lack of consensus does not imply that some of them are irrational. However, arguably, if I am one of the jurors and I right now believe that the defendant is guilty, I should not change my mind unless I receive new information, notice a mistake in my previous reasonings, or something like that; that is, perhaps [NOSWITCH] does apply in a nonderivative manner to such cases.

But let us examine the plausibility of [NOSWITCH] in such cases more carefully. Let us assume that certain epistemic dispositions determine for a subject which permissible attitudes they will take. Suppose rational subjects can differ in how epistemically anxious they are, so that more anxious agents need more evidence to form a belief than less anxious agents.¹⁹ We assume

¹⁸Rosen (2001), for instance, argues for this possibility, while White (2005) argues against it. The jury example below is from Rosen (2001).

¹⁹Of course, the disposition need not be context-insensitive or interest-invariant. But it will keep our lives simpler to ignore these complications.

that these dispositions may change throughout the course of one’s epistemic life. But now suppose you had one of the rational dispositions D . However, after a while you become more anxious and you now have disposition D^* . Let us assume that D^* is still within the range of permissible dispositions. Could the mere fact that your disposition changed be irrational? It is hard to argue that it could. It seems that if I and You can be both equally rational while holding these different attitudes to the same proposition given the same evidence, one should grant the same courtesy to I_{t_0} and I_{t_1} when ‘we’ disagree.

Now suppose that it is also possible to form beliefs that deviate from the disposition. It seems that whether a belief that I form deviates from my anxiety disposition should be irrelevant to whether it is a rational belief. The fact that an anxious person like me has a more confident moment should not on its own make me irrational if the attitude I form due to this sudden burst of confidence does not lead me away from the permissible range of attitudes.²⁰ Given that, at least as far as we have stipulated, the combination of my dispositions and my “deviant” behaviour never leads me astray, it seems that there could be nothing wrong with my pattern of forming epistemic attitudes. If it is (roughly) sufficient for a disposition to be rational that, from justified attitudes, it leads us only to justified attitudes, the pattern of dispositions and deviations in such a subject is also a rational disposition.

This argument is by no means decisive. One could argue, for instance, that even when multiple permissible attitudes are available, the rational subject must take her attitude to be required by her evidence. This would result in a fragmented life for the rational, philosophically sophisticated believer; such a subject would have an attitude to the evidence that she knows, upon reflection, to be unwarranted. But still it is not obviously incoherent.²¹ However, I will not pursue the issue further here, and just rest content to have established that arguing for independent persistence requirements on beliefs will be an uphill battle. The best strategy for establishing such requirements for intention focuses on features of intentions and intention formation that

²⁰One might argue that it would be a further rational requirement that my pattern of belief formation should not arbitrarily vary from proposition to proposition. Perhaps if I form the belief that \mathbf{p} based on evidence E , I should also form the belief that \mathbf{p}^* based on E^* if there is no relevant difference between the epistemic relation that \mathbf{p} bears to E and \mathbf{p}^* bears to E^* . I doubt that there is such a requirement, but if it were valid, it would still be a requirement among synchronic beliefs. This kind of requirement would not imply [NOSWITCH], but only would require that such switch would be accompanied by similar change in attitudes held based on similarly “strong” evidence.

²¹White (2005) considers some versions of this possibility, and argue against them. I am sympathetic to what White says there, but this issue lies beyond the scope of this paper.

are not shared by beliefs.

Persistence of Intention

Some features unique to intention might seem to give rise to some kind of persistence requirement:

1. In the practical realm, ‘ties’ are more widespread, and it is important to ‘break’ these ties.

Often when we pick from a restaurant buffet, get something from our pantry, choose dates to travel, pick a time to go to the bookstore, etc., we need to pick among indifferent alternatives.²² It is however essential in such situations that we ‘break the tie’; we should not share the fate of Buridan’s ass when we face alternatives that are all equally good and better than inaction. Forming an intention in such a case arguably settles which of the indifferent ends one has adopted. Arguably, then, it is a constitutive condition,²³ or a rational requirement, of settling in this way that the intention persists.

2. Cases in which an agent faces incommensurable alternatives are widespread in the practical realm and they require that the agent choose an alternative without either being better than the other and without the options being equally good.

From deciding between going to a movie or a restaurant to the decisions about which career to choose, we face choices between alternatives that do not seem to be comparable.²⁴ Forming an intention in such cases arguably settles which of the incommensurable ends one has adopted. Arguably, then, it is a constitutive condition, or a rational requirement, of settling in this way that the intention persists.

²²Ullmann-Margalit and Morgenbesser (1977) famously argued for the ubiquity of cases of ‘picking’ (selection under indifference). Some of the cases of what they call ‘second-order picking’ are cases of incommensurability and will be discussed separately.

²³Phrases such as ‘constitutive condition’ are often slippery and difficult to provide a precise definition for. My contention is that no matter how one tries to make this notion precise, one will not be able to escape the arguments below against persistence requirements.

²⁴How to understand these cases more precisely is a matter of much debate. For some suggestions, see Chang (2002), Griffin (1997), and Gert (2004). I’ll try to remain somewhat neutral among those different ways of understanding the phenomena, but I place some ordering constraints below that are relevant to this debate.

3. Forming a future-directed intention, by its very nature, settles among alternatives previously left open. There is no other point in forming a future-directed intention. It is unclear how future-directed intentions could have such a function unless it was a rational requirement that they persist.

Of course, forming a belief also settles an issue that was open before. But it is not epistemically rational to form a belief simply because we need to settle on a question; rational agents form a belief because the evidence rationally compels them (or at least permits them) to form such a belief.

(3) really generalizes the points made in (1) and (2). But if what we said above is correct, ties and incommensurable alternatives are the best cases for intention specific requirements. In fact, insofar as it is true that rational belief is ‘compelled’ by evidence, intentions can also be rationally ‘compelled’ by the reasons to form them. Perhaps *when* we form the intention is not rationally compelled, but we know that we have independent reasons in many cases to form intentions prior to the time of action as a necessary condition of realizing something we deem good or something we believe we have reason to pursue. It is in the cases of ties and incommensurability that we can make the best case for intention specific requirements.

One might argue that other ends we have justify certain other intention requirements. Proposals like this have been made with respect to the aims of coordination (Bratman 1999), self-governance (Bratman 2009), and resisting temptation (Holton 2004). But these constitute external justifications and we leave them for another occasion.²⁵

Ties

Let us start with the following vignette:²⁶

There will soon be a fork on the road. Both branches take me to my destination in the same amount of time, without any significant difference in the landscape, gas costs, etc. I form the intention now that I will take the left branch.

A few minutes pass and I now reach the fork. No one doubts, of course, that if I carry out my intention I acted rationally, *ceteris paribus*. But suppose that my driving comes to an unexpected conclusion. Let us take a look at those two alternative endings for our gripping vignette:

²⁵I argue that these justifications also fail in Tenenbaum (in preparation).

²⁶Taken, of course, from (Bratman 1999).

- (a) The intention slips my mind, and I turn right.
- (b) I still remember having formed the intention to turn left, but I abandon it for no particular reason, and turn right.

Are these alternative endings instances of irrationality? Given that I successfully and non-accidentally did something that I was (rightly) indifferent towards compared to the option of taking the left branch, it seems hard to explain why this would be irrational. It certainly seems that I did not violate any instrumental requirements, since the path I took was just as good in bringing about my ends or satisfying my preferences as the one I forsook. However, one could argue that I did violate a requirement; after all, what would have been the point of forming a future-directed intention in the first place, if it turned out that it made no difference to what it was rational for me to do? We can even think in terms of a dilemma. Either (i) there was a reason to form my intention and thus there would be reason to carry it out;²⁷ or (ii) there is no reason to form my intention and thus it was irrational to form it. The fact that we can come up with such a dilemma, arguably, suggests that the combination of forming the intention and not carrying it out is a case of irrationality.

However, this argument suffers from an ambiguity. The first horn of the dilemma is true only if there is *decisive* reason to form the intention. That is, even granting that a reason to intend to φ must also be a reason to φ , it would follow that I now must φ only if the reason in the case was decisive, only if there was not just as good reason to do something else. But, *ex hypothesis*, there is just as good reason to turn right. On the other hand, the second horn of the dilemma is only true if there was no *sufficient* reason to φ and thus no sufficient reason to form the intention to φ . Even if there was no decisive reason to form the intention (that is, even if there was also sufficient reason to form another, incompatible, intention), it would not be irrational to form the intention as long as there was sufficient reason to form it. Yet, the dilemma does seem to have an intuitive appeal. How could intentions settle what we do if it were not for the fact that they are going to persist? And if a state does not settle what we do, how can it count as an intention?

It is correct to say that it is of the nature of intentions that agents who form an intention to φ are *not* just as likely to φ as not to φ . Moreover it is

²⁷Of course my opponent can readily grant here that this is not generally true. Perhaps in some cases there are reasons to form an intention to φ when there is no reason to φ , as in the Toxin puzzle. There are complicated issues here, but all parties to the debate would accept that our vignette is not a special case in which the reasons to form the intention are completely independent from the reasons to perform the action. See Kavka (1983).

exactly because intentions tend to persist that they can serve the function that they do. In our earlier case of choosing among travel destinations, it was important to start planning for a trip, since I had decisive reason to go somewhere. Moreover, even if I need not do anything right now, it might be worth deliberating now as it might be costly later.²⁸ And what we said earlier about preservation of justification is correct here too. Whatever justification I had in forming the judgment that I should φ (or that φ is one of the things that I am permitted to do at a certain choice point) is preserved together with the judgment. *Ceteris paribus* my justification for forming the intention to φ is preserved whenever my intention to φ is preserved. None of this, however, implies a requirement of any kind. If the justification for my preserved intention *permitted* the intention, but did not *require* it, then I am no more required to φ now than I was required to form the intention to φ then. We can agree that the function of a future-directed intention is to settle an issue and, as it were, set the agent in the straight path to act from this intention, without needing to postulate any normative requirements that intentions generate. Barring interferences (including forgetfulness), intentions persist till they are acted upon. This fact is enough to explain how they can serve their ‘settling function’. In our original vignette, the point of forming the intention was, no doubt, that it would save you from having to deliberate while impatient drivers behind you wait for you to make a decision. Given that you were justified in forming the intention to turn left earlier, you are still justified in acting on this intention when you reach the fork. Future directed intentions serve their settling function only because in the normal course of events at the time of action they become intentions in action. But this does not imply that my interfering with this normal course of events is (even presumably) irrational.

What we do in tie cases is a matter of indifference; it is no surprise then that future-directed intentions which simply settle what we will do in such cases do not generate any persistence requirements, or at least none that on their own will have much intuitive pull. Cases of incommensurability might be better to illustrate a supposed special role of future directed intention, since these are, by definition, cases in which one is *not* indifferent between various courses of action and in which one can invest a great deal into the pursuit of one alternative without altering the comparative relation between two choices. If my buying a nonrefundable ticket to Paris sufficed to making going to Paris superior to going to London, given the higher cost of the

²⁸Ferrero describes this role of intention as a division of deliberative labour. See Ferrero (2010).

London alternative, the same cannot be said for choice between, say, giving a large sum of money to my favourite charity or spending the money on a very expensive family reunion. These alternatives might still be on a par even after my travel agent offers a 20% discount on the family reunion package.²⁹ In such cases we might think that there is a more important role to be played by the stability of a future-directed intention and this might be, consequently, a better context for the justification of a persistence requirement. So we should see whether cases of incommensurability can provide better grounds for persistence requirements.

Incommensurability

Raz's famous proposal defines two alternatives A and B as incommensurable when it is neither the case that A is better than B nor that B is better than A, nor that A and B are of equal value.³⁰ But it is also possible to talk about values being incommensurable, if, for instance, tradeoffs between different kinds of value are not always comparable.³¹ So we can say that a specific choice between watching "Tokyo Story" and going to dinner at Chez Panisse is incommensurable, but we can also talk more generally about the value of aesthetic appreciation and gastronomic enjoyment being incommensurable. Neither notion of incommensurability says much about how an agent ought to act in situations involving incommensurability. A natural way to respond would be to say that all choices in such situations are permissible. Now if we focus on the second notion of incommensurability, this is not quite correct. After all, there could be a situation in which a choice is inferior in all aspects. So even if aesthetic appreciation and gastronomic enjoyment are incommensurable, if my choice is between watching "Tokyo Story" and having dinner at Chez Panisse, on the one hand, and watching "Battlefield Earth" and eating at Subway on the other, it will be arguably the case that I must choose the former. However, it is still correct that there is a wide array of choice situations and possible alternatives in which choices involve a plurality of values and ends, and it might be that there will be whole range of

²⁹I using Chang's notion of 'being on a par' (Chang 2002), but this point, of course, does not depend on accepting Chang's views on incommensurability.

³⁰Raz (1986, p. 342). Of course, one could define the notion similarly in terms of preferences.

³¹For an example of the latter use, see Wiggins (1978). Chang (1997) distinguishes between incomparability and incommensurability along similar lines and provides a very useful guide to the various ways of understanding 'incommensurability' and similar notions.

permissible choices among alternatives that are not equal. It will be easier in our discussion to proceed with a more precise representation of incommensurability, but the argument will not depend on this representation.³² First I will talk about ‘ranking (or value) within an evaluative dimension’ rather than a ranking as a manifestation of a certain value. This admittedly cumbersome language leaves room for the possibility that commensurability does not go hand in hand with instantiation of the same values. Not all cases in which tradeoffs are not comparable can be plausibly described as cases in which we realize different kinds of values. If I am lucky enough to have two friends, various choice situations involving the two of them will have a similar structure to the choice situations in which I could realize two kinds of values. But it would be implausible to describe, say, helping Jane train for the marathon and helping Jay train for the marathon as the realization of two different kinds of values.³³ The notion of evaluative dimension is supposed to be more eclectic; we can stipulate that all the things (options, instances of goods, object of preference, or whatever bearer of value one chooses) in the same evaluative dimension are weakly ordered. We can then represent an option by an ordered set of values across evaluative dimensions. So if we take the relevant evaluative dimension to be something like ‘aesthetic appreciation’ and ‘gastronomic enjoyment’ in our options above, we could attribute to our options above of [Tokyo Story, Chez Panisse] and [Battlefield Earth, Subway] the pairs of values, respectively, [1,1] and [0,0].³⁴ We said above that in this case, the choice of Tokyo Story and Chez Panisse is the right one given the obvious aesthetic and gastronomic stipulations. This suggests the following, extremely plausible Basic Ordering Principle for the ordered sets:³⁵

[BOP] if $x_a > y_a$ and $x_b > y_b$, then choosing $[y_a, y_b]$ is not rationally permitted.

Such a principle still leaves many options and would allow for various permissible options that are not incompatible. To see how weak the principle is it is worth briefly comparing it with a principle that some attribute to Sidgwick. Sidgwick is often thought to have considered well-being and morality to be

³²This proposal bears some similarities to the representation of value relations in Rabinowicz (2008).

³³This is similar to a point made by Thomas Hurka regarding rational regret. See Hurka (1996).

³⁴Since we introduced no basis for cardinality, these numbers should be understood as ordinals.

³⁵For simplicity sake, I will look only at the two dimensions case.

in some similar way incommensurable. Under some interpretations, however, Sidgwick provides a much stronger ordering principle; a principle that requires that I always choose either the best option from the moral point of view or the best option from the point of view of my own well-being. Let us call this principle, in a somewhat unsympathetic fashion, Sidgwick's Insane Principle:³⁶

[SIP] It is permissible to choose $[x_m, x_w]$ only if EITHER (for every other option $[y_m, y_w]$, $x_m \geq y_m$) OR (for every other option $[y_m, y_w]$, $x_w \geq y_w$).

Under certain interpretations of Sidgwick, he says that a rational agent either maximizes the overall good or her own good. This would mean that if I earn a million dollars I could either keep it all to myself (given certain admittedly strong assumptions about personal good) or give it all away to charity. But, for instance, the compromise in which I keep \$100 for myself to buy new headphones and give the rest to Oxfam is deemed irrational. [BOP], on the other hand, only rules out the cases in which the use of the money is inefficient (if, for instance, I pay \$200 for the headphones when I could have just as easily paid \$100, while sending my Oxfam donation via an expensive courier, when I could just as easily donated online). At first the focus on incommensurability might seem not to introduce anything different from the case of ties. Certainly, if we focus in a case like the simple case below, we have no reason to find any more justification for an intention specific requirement than when we looked at the cases of ties:

(SIMPLE)

I can watch either "Better Movie" or "Movie" at "Town Theatre". My other option is to go to "Fancy Restaurant". I don't prefer "Fancy Restaurant" over either "Better Movie" or "Movie" or vice versa. However, I form the intention to go see "Better Movie". I start driving through "Road to Everything" but just before I need to turn to "Movie Road", I abandon³⁷ my intention to go to the movie, and take "Restaurant Road" instead.

³⁶This is unsympathetic labeling, but, in my defense, I am in no way confident that Sidgwick actually holds it. Gert (2007) attributes it to him in the course of attempting to refute the view. I do think, however, that much of what Sidgwick says is compatible with weaker principles of choices.

³⁷I mean to use 'Abandon' here without any connotation that I *do* something such as reconsider, actively change my mind, etc. Simply dropping the intention counts as 'abandoning' for my purposes here.

If we were persuaded by the argument that I violated no requirement when I turned right in our above vignette, we should be equally convinced that I have violated no requirements in (SIMPLE). However, if the only constraint on permissible choices we have in these cases is [BOP], it seems that we need to allow much more radical changes of mind if we do not introduce intention specific requirements. Let us look at the following better case:

(BETTER)

Larry is deciding between being a professional footballer or a stay-at-home Dad. In order to become a professional footballer, he must buy a \$200 ball and net set. If he wants to be a stay-at-home Dad, he needs to buy the “How to Be a Stay-At-Home Dad” DVD for \$200. Larry forms the intention to become a professional footballer, goes to the store, and buys the ball and net set. The next day he abandons his intention, calls the Barcelona manager, and says that he no longer wishes to be in the team as he is now a stay-at-home Dad.

This seems to be a case in which it is plausible to say that this decision is one in which Larry finds the two options (stay-at-home Dad; Professional Footballer) incommensurable, and that a difference of \$200 dollars in the cost of each alternative would not suddenly make one of the options better than the other. So taken in isolation, either choice seems permissible. That is, choosing to be a professional footballer at the first juncture is obviously permissible. But also it would have been permissible to choose to become a stay-at-home Dad if the cost of this option was \$200 higher. Yet it seems that something went awry here. Larry seems to have done something foolish; after all, he just wasted \$200. Yet, [BOP] on its own cannot condemn him. There are two evaluative dimensions relevant for Larry’s choice; we can call them the “professional” dimension and the “family” dimension. Now if A = [stay-at-home Dad], B = [stay-at-home Dad and loss of \$200] and C = [Professional Footballer] and D = [Professional Footballer and loss of \$200], then A, B, C, and D are assigned the following set of values: A = [0, 2], B = [0, 1], C = [2,0], D= [1,0]. But obviously neither Larry’s earlier nor Larry’s later choice violate [BOP].

However, I’ll argue that appearances here are deceiving. We have good reason to suspect that the intuition that Larry is irrational in this case is misleading.³⁸

³⁸Considering a similar example, Broome (2001) argues that an agent in Larry’s situation is rational as long as he *repudiates* his previous intention. But if minimalism is right,

Challenging the Irrationality Intuition

Let us start by looking at a variation of Better Case:

(MODIFIED BETTER)

Larry faces the same situation and forms the same intention and buys the same ball and net set. Larry then thinks he gets a call from his team manager (Gerardo) and thinks he hears Gerardo say: “You’re dropped from the squad”. He then abandons his intention to be a professional footballer (as he now thinks that this career is no longer available to him). Before he even has time to consider what he’ll do instead, the phone rings again. It’s his cousin Heraldo. Heraldo tells him he was the one who called him and that what he said was: “I got ear drops in Riyadh”. Without considering the matter further, Larry immediately forms the intention to become a stay-at-home Dad.

It seems hard to distinguish (MODIFIED BETTER) from (BETTER) in terms either of the intuitions they elicit regarding Larry’s rationality or the well-groundedness of these intuitions. Yet (MODIFIED BETTER) could not possibly be a case of failing a requirement not to abandon an intention. Assuming that Larry was justified in forming the belief that he had been dropped from the squad, then his abandoning his intention was perfectly rational; in fact, keeping the intention would be irrational in light of this belief. He also could be under no requirement to revise the intention to be a professional footballer; given our stipulations, it would be no less rational to form the intention to become a stay-at-home Dad. In fact, we can see a further problem with postulating persistence requirements if we think about a small variation to this case:

(FURTHER MODIFIED)

Same as (MODIFIED BETTER) except that Larry does form the intention to become a stay-at-home Dad between the first and the second phone call.

Would Larry now be under a requirement to continue to intend to be a stay-at-home Dad after he learns about his mistake? It would be counterintuitive to say “yes”. Given that this was what he was planning to do before the

the agent is rational as long as she is not engaged in incompatible pursuits; that is, as long as Larry is no longer pursuing the footballing career, he is rational; no ‘repudiating’ act is necessary.

confusion, wouldn't it be rational for Larry to go back to his original intention? And wouldn't the rationality of this course of action be independent of the content of further deliberation or reconsideration or even of whether any further deliberation or reconsideration occurs? But it is hard to see how an advocate of persistence requirements can accept this conclusion. It is true that Larry does get new information, so his updated belief state might be seen as 'covered' by the *ceteris paribus* clause of the correct persistence requirement. However the new information does not make a difference to any of the relevant values or evaluations in Larry's situation. It is still true that neither choice of profession is better than the other, and it is still true that they are not of equal value. This suggests that something might have gone wrong with our apparently intuitive judgment that Larry's actions were irrational at any point.

Pattern and Instance

Let us take a different perspective on Larry's choice in (BETTER). Suppose that Larry did change his mind and became a stay-at-home Dad, and everything went as he expected. He was happy with such a life, his assessment of the value of such a life was overall correct,³⁹ and there were no surprise, twists, etc. that would be relevant for our later assessment of Larry's life. Larry is also (correctly) aware that had he chosen the path of a professional footballer, he would also have realized an important value and another end of his; he neither regrets his decision from this vantage point nor thinks that his life would have been worse (or better) had he chosen to be a professional footballer. Does it still seem that Larry originally acted irrationally? Should we say that, fortunately, everything turned out well, but that his decision to become a stay-at-home Dad was irrational? It seems that given that his life, predictably, did not turn out to be worse than had he held on to his intention to become a professional footballer, his choices were all permissible. Of course, often irrationality is harmless and sometimes it might even pay off. But, in general, these are cases in which circumstances luckily conspire to make it the case that irrationality does not lead us astray. But in our case. Larry's judgments are, *ex hypothesis* all correct, nothing unexpected happens, etc.

This vantage point on Larry's choice suggests that an inclination to judge Larry irrational lies somewhere else. Notice that were Larry to continue

³⁹I am adding these claims about the correctness of his judgment to avoid complications about assessing the rationality of an agent in light of various errors. Arguably, none of these stipulations makes any difference.

changing his mind in this way, the results could be disastrous. Suppose that every time Larry switches from one intention to another he gives away the book or the ball and net set to avoid clutter. Thus every time, he changes his mind it costs him \$200. If he keeps doing this, he will be much impoverished and yet his life options would be the same. Doubtless, if Larry keeps changing his mind in this manner, he has behaved in an irrational fashion. It is hard not to project onto Larry this kind of fickle disposition that is sure to end in disaster. However, the irrationality of proceeding this way need not appeal to any intention persistence requirement; it is simply a consequence of the rather well-known fact that for all (or at least most) of us, having some minimal financial resources is necessary for the realization of our ends or of much that is valuable; given this fact, we are required not to engage in actions that will lead to poverty without realizing any other value or end.

One could try to defend an intention persistence requirement exactly on the grounds that multiple instances of such changes of mind are irrational. If multiple instances are irrational, and all these instances are alike, then there should also be a requirement against any instance of failure of persistence. But this argument is flawed. It does not follow from the fact that it would be disastrous to accept every drink offered to me at a party (since I would get drunk and make a fool of myself) that I should turn down every single one of them. We cannot derive a requirement never to φ from a requirement not to φ too many times.⁴⁰ Still, it is hard to avoid projecting into Larry this kind of disposition. So given that the judgment of irrationality from the vantage point of later in Larry's life does not seem particularly intuitive and in other ways problematic, it is plausible to conclude that our temptation to deliver a verdict of irrationality in (BETTER) stems from this projection.

Further Constraints

However, there is a way to read (BETTER) in which Larry is in fact violating a constraint, but not an intention persistence requirement. In particular, (BOP) does not capture all the constraints that choice among incommensurable evaluative options should answer to. Let us look at a modified version of (SIMPLE):

(SIMPLE MODIFIED)

I jotted down a series of notes on the movies showing tonight

⁴⁰I argue in more detail against views that attempt to derive a requirement prohibiting specific actions from the irrationality of a certain pattern of activity in Tenenbaum (2010, in preparation), Tenenbaum and Raffman (2012).

and instructions on how to get to the “Town Theatre” and how to get to “Fancy Restaurant”. I can barely understand them now given my poor handwriting. I can’t tell from my notes whether the movie showing at the “Town Theatre” is “Movie” or “Better Movie”. I understand just enough of my notes to conclude the following. Road A will take me to the restaurant if and only if “Better Movie” is showing. Road B will take me to the restaurant if and only if “Movie” is showing. Road C takes me to the restaurant. Road D takes me to the movie, no matter what it is. I take Road A.

There seems to be something incoherent in my attitudes in (SIMPLE MODIFIED). I could have chosen just to go the movie no matter which one was showing, and I could have chosen just to go to the restaurant. I could have made it the case that I would go the movies, if “Better Movie” were showing and otherwise to the restaurant, but instead, I chose to go the movies if “Movie” was showing, but to the restaurant if “Better Movie” were showing. In other words, the only overall ranking of the alternatives that could be read from my choosing in this manner is, from higher to lower, {Movie, Fancy Restaurant, Better Movie}. On pairwise choices, I rank “Better Movie” higher than “Movie”; it seems that the introduction of another option (Fancy Restaurant), even if incomparable with my movie options, should not lead to choice disposition that invert this ranking.⁴¹ Although such situations are rare, the incoherent attitudes are arguably present as choice dispositions even in the absence of such option. That is, my attitudes seem to be incoherent if I would choose to go “Movie” over “Fancy Restaurant”, but I if were to learn that “Better Movie” were showing then I would choose to go the restaurant instead.⁴²

Yet this combination of attitudes does not violate [BOP]; a further ordering constraint is needed:

[FOC] If $[x_a, x_b] \geq [y_a, y_b]$ and $z_a \geq x_a$ and $z_b \geq x_b$, then $[z_a, z_b] \geq [y_a, y_b]$

⁴¹Of course, we need only Road A and Road B to generate the problem. In fact, the choice of Road A over Road B is already incompatible with the axioms of classical decision theory as long as one keeps the probabilities of ending up at the restaurant constant (assuming that I rank “Better Movie” above “Movie”). But appeal to the axioms of classical decision theory in this context is obviously controversial.

⁴²Since arguably even the choice between two movies is a choice along incommensurable evaluative dimension, it might be easier to see the incoherence if we think of “Better Movie” just as a better print of the same movie.

Does Larry violate [FOC] in our original case? Since we did not specify what his counterfactual attitudes were in describing (BETTER), we left it open whether his attitudes were, at the initial point, that he would choose to be a professional footballer over being a stay-at-home Dad, but he would choose the latter if it would cost him an extra \$200 to be a stay-at-home Dad. However, [FOC] is a synchronic requirement on choices or choice dispositions. So there is a way of reading (BETTER) that makes Larry incoherent, but even in such a reading Larry's incoherence is not due to the violation of a (diachronic) intention persistence requirement.

It is important to note that for my purposes, rejecting [FOC] would only make my case stronger. For if you do not accept [FOC], there is no coherent way that you can argue for intention persistence from Larry's supposed intuitive irrationality in (BETTER). If you reject [FOC], you have already accepted the rational acceptability of choice dispositions in which Larry would choose to become a stay-at-home Dad if this will cost him an extra \$200. But this is exactly what happens in (BETTER), so you cannot now say that a requirement of rationality forbids him from making this choice.

To sum up, there are a number of difficulties in trying to hold on to the view that a persistence requirement is violated in (BETTER). Moreover there are a number of irrational attitudes in the neighbourhood that might account for whatever inclination one has to judge that Larry is being irrational in this scenario. I hope that these considerations are enough to lay to rest the claim that cases such as (BETTER) suffice to establish the cogency of intention persistence requirements.

It is worth noting that versions of [PERSIST] that count *forgetting* an FDI as a violation of [PERSIST] will have counterintuitive consequences, at least if one thinks that there is a reason to be rational. We said above that, in the case of belief, we should reject the idea that forgetting is a form of irrationality. Some of the reasons we give there do not apply to the case of future-directed intention; FDIs do not turn useless with the passage of time, at least not until they are executed, in any reliable way. But we can see the implausibility of this view by looking at the following case. Suppose that I'll need to decide 50 times on my trip whether to turn left or right, and that in each case they're equally good routes to my final destination. I suffer from decision exhaustion and if I make too many decisions while driving, I'll become tired and irritable, and have an unpleasant drive. But now I am relaxed, so it's a good time to make decisions. I realize that I'll not remember every single intention I have formed, but I'll remember enough (say 40), and 10 decisions will not change my mood in any significant way. Now on a view that forgetting the intention would count as irrationality, assuming we have

reason to avoid irrationality, there would be some reason for me not to make the decisions earlier. This already seems implausible, but reasons are cheap, and one might think that there is *some* reason not to make the decisions now. But if these reasons have some weight, we should be able to find a case in which the balance of reasons is close enough, so that the reason not to be irrational will tip the balance in favour of not forming the intentions now, and, under certain assumptions about how these reasons add up, the more instances of such forgetting, the easier will be for them to tip the balance in favour of not making the decisions early. But this seems a rather implausible consequence; the fact that one will be going to forget too many intentions cannot by itself be an essential part of the explanation of why one decided against making decisions ahead of time in such a situation.

Means-Ends Coherence

It is very plausible that some principle of reason transmission from means to ends is valid, and I do not want to deny that there is a normative principle of means ends coherence. However, I will argue that no plausible version of a means-ends coherence principle threatens the position I defined as Minimalism about intentions; I want to show that no underivative principle of means-ends coherence applies to future-directed intentions rather than intentions in action. Let us first happily grant the truth of a version of Raz's Facilitative Principle (Raz 2005):

[FACILITATIVE] If there is reason to φ , and one can only φ if one ψ -s, then one has reason to ψ .⁴³

Raz puts forward [FACILITATIVE] in the course of arguing that there is no general requirement of instrumental rationality, and later authors have defended the more radical view that there are no requirements of practical coherence (Kolodny 2008). Trying to argue for minimalism on the basis of these claims would be overkill; we would show that there are no intention specific rational requirements by showing that there are no rational requirements.⁴⁴ But this is not necessary. For once we accept [FACILITATIVE], we can even more happily grant a coherence requirement as follows:

⁴³This is, of course, not the same formulation as the one provided by Raz. This is a rough formulation of a transmission principle; in particular, various complications arise in cases in which I will not or should not ψ . The precise formulation does not matter for my purposes, as long as the correct principle does not make ineliminable reference to (future-directed) intentions.

⁴⁴Or by showing that no rational requirements are normative.

[FACILITATIVE_C] If A judges that there is conclusive reason for her to φ , and one can only φ if one ψ , then she should judge that she has reason to ψ .⁴⁵

But the minimalist can also accept that there is a more proper principle of instrumental rationality that applies to intention in action or just acting intentionally or having an end:

[INSTRUMENTAL_A] If A is φ -ing intentionally and she (judges that she) cannot φ unless she ψ -s then she is rationally required to ψ .

One might think that [INSTRUMENTAL_A] is always trivially satisfied; after all, if I did not take the necessary means, I also did not perform the action. However, this is not quite true, at least not for the relevant action identified by [INSTRUMENTAL_A]. Even if I cannot, for instance, shop for bread without lining up at the checkout lane, I can be *shopping for bread* without lining up at the checkout lane; it is only that if I never line up, it will never be true that I successfully *shopped for bread*.⁴⁶ Perhaps there are some actions in which the failure to take necessary means makes it the case that we do not even count as φ -ing. If I go to the wrong store, without any cash in my pocket, etc. perhaps it is not even true that I am shopping for bread. But even in these cases I am still engaged in a number of actions that count as trying to φ . So if we want to allow for the possibility that in cases of massive error I will not even count as φ -ing, but I will still be subject to a norm of instrumental coherence, we can modify our principle accordingly:

[INSTRUMENTAL_{A*}] If A is φ -ing intentionally or trying to φ and A judges that she cannot φ unless she ψ -s, then she is rationally required to ψ .

Is there a *further* requirement of rationality not covered by [FACILITATIVE_C], [INSTRUMENTAL_{A*}], or [INSTRUMENTAL_A]? Should we also postulate something like the following principle?

[INSTRUMENTAL_I] If A has (a future-directed intention) to φ_{later} , and (she believes that) she can only φ_{later} by ψ -ing, she must also intend to ψ .

⁴⁵As stated, the principle reads more naturally as a narrow scope principle, but I am not committed to understanding the principle this way. Whether “should” takes wide or narrow scope is not relevant for my purposes.

⁴⁶Michael Thompson has recently emphasized the importance of the progressive aspect in understanding action and the point that the truth of the progressive does not depend on success in one’s endeavours. See especially Thompson (2008).

But here we can see that if there is no requirement for intention persistence, [INSTRUMENTAL_I] will also fail to be a genuine requirement. Let us go back to our graphic above with our travel decisions. The same thing we said about intention persistence applies here. Until t_2 , any instrumental requirement I am under will be a consequence of [FACILITATIVE_C]. The same thing can be said about the requirement I am under after $t_3 + i$ and before t_4 . Finally after t_4 the requirements I am under will be a consequence also of [INSTRUMENTAL_{A*}], or [INSTRUMENTAL_A]. Just as in the cases of intention persistence requirements, the best chance for [INSTRUMENTAL_I] doing any work lies in the evaluation of the rationality of the agent between t_2 and t_3 . But if there is no requirement that the intention persist, why should there be a requirement to pursue the necessary means? If it would have been rational for me to abandon the intention without any deliberation, why wouldn't it be rational for me not to pursue the means? Both will have the exact same result; namely, that I do not go to Paris on vacation. But since I judge London and Paris to be equally as good, as long as I make sure that I do go to one of these destinations, my actions seem no less rational.

In fact, it is hard to distinguish between knowingly not taking the necessary means to a future-directed intention and abandoning the intention, if I am doing nothing else in order to realize this intention. But even if we can say in some cases that the intention continued despite the fact that one did not take the necessary means, then, given that abandoning the intention is a rational option for me in these circumstances, it seems that the most I am required to do is *either take the means or abandon the intention*. Since my not having taken the means is now irreversible, I can now detach the second disjunct and conclude that I must abandon the intention.⁴⁷ Moreover, the disjunctive requirement is a straightforward consequence of [FACILITATIVE_C]. An unrealizable intention can only generate "costs".⁴⁸

One might object that a future-directed intention is not simply idle until we have the opportunity to take necessary means to its realization. Future-directed intentions often warrant the pursuit of actions that are not strictly necessary for their fulfillment. So, perhaps I have taken French lessons, which were not strictly necessary for my trip to Paris, but which would improve my experience. But now it seems that if I have done all those

⁴⁷I am assuming that irreversibility here validates detachment. This is somewhat controversial, but it would take us beyond the scope of this paper to try to defend it.

⁴⁸There could be deviant cases; it seems easy to imagine cases like the toxin puzzle in which I am highly rewarded for having an unrealizable intention. But whatever one wants to say about such cases, one cannot use them to justify an intention persistence requirement.

things and fail to take the necessary means to go to Paris, I failed in some requirement of rationality. However, this objection simply takes us back to issues that we discussed earlier. If taking French does change the value or preference relations (or my ends), then I might now be under a requirement to take the necessary means to go to Paris, but this will be accounted for by [FACILITATIVE_C]. If not, our other principles cannot generate any requirement but if our discussion of ties and incommensurability above is correct, there is no valid requirement to be generated. Similarly, I could fail to take means that were not strictly necessary but made the realization of my intention much easier. I could, for instance, fail to buy cheap tickets to Paris, and still have my intention to go to France. But, again, if the opportunity to buy the cheap tickets would not have put me under a requirement to buy the tickets independently of the intention, we have no reason to think that merely having the intention could create such a requirement given that there is also no requirement that my intention persist. Of course, if I *predict* or am reasonably confident that I will end up going to Paris, then I have a reason to buy the ticket. But this is also a straightforward consequence of [FACILITATIVE_C].

External Grounds

As we said above, this is only a partial defense of Minimalism. After all, we have not examined the possibility of external grounds to accept intention specific requirements, and philosophers sympathetic to such requirements have greatly emphasized these kinds of external grounds. Exactly because of their importance, these external grounds deserve separate treatment, and any serious discussion of them must be left for another occasion.⁴⁹ Here I can only mention very briefly why I think they would not work. There are basic two kinds of justifications for such requirements: one that depends on certain ends that we have (coordination, self-governance, etc.), and another that depends on certain failings of human nature (succumbing to temptation, susceptibility to biases, etc.). The first kind of justification often appeals to an end that human beings undoubtedly have, but, of course, none of these ends are the *only* ends human beings have or even ends that have lexical priority over other ends. It is thus hard to see how they could generate *requirements* rather than just certain reasons in favour of certain actions.

The second kind of justification has a peculiar structure: these requirements would tell us to reason in a certain way when we would otherwise

⁴⁹I argue against these kinds of grounds in (Tenenbaum in preparation).

reason in a different way simply because we are susceptible to certain kinds of rational failing. Of course, I know I am susceptible to miscalculation, and when I suspect I might be misled I must try to ensure as well as I can that this does not happen; if I am generally prone to addition errors, and I have little time to make a decision, I might prefer to trust a friend's calculation instead of my own. This is not an independent requirement at all; it is a rather direct consequence of an instrumental requirement in pursuing a certain end. An independent requirement would demand that I follow it even if I do *not* independently judge that I am more likely to hit the target by not trusting my assessment of the 'first order' evidence. It seems that, on this view, you can describe the agent facing temptation as follows: such an agent understands that she is under a rational requirement not to reconsider her initial intention. She also understands that she should not reconsider because were she to reconsider she might reject the intention on grounds that she (at least now) does not consider to be good grounds to reject the intention. And she does that not because she understands that she is more likely to make a mistake if she reasons in a different manner, or because she values commitment, or for any other similar reason. This kind of position seems to stretch the limits of coherent agency.⁵⁰ But of course here I am presenting something close to a caricature of these intricate and sophisticated positions; settling these issues require much further discussion.⁵¹

References

- Anscombe, G. E. M.: 2000, *Intention*, 2 edn, Harvard University Press.
- Bratman, M.: 1999, *Intention, plans, and practical reason*, Cambridge University Press, New York.
- Bratman, M.: 2009, Intention, practical rationality, and self-governance, *Ethics* **119**(3), 411–443.
- Bratman, M.: 2012, Time, rationality, and self-governance, *Philosophical Issues* **22**(1), 73–88.

⁵⁰See (Paul 2011) for a similar criticism of Holton's views in (Holton 2009).

⁵¹I would like to thank Chrisoula Andreou, Michael Bratman, John Broome, Luca Ferrero, Jennifer Nagel, Sarah Paul, Wlodek Rabinowicz, Benjamin Wald, and audiences at the Center for the Study of Mind in Nature (Oslo), Oxford Area Action Theory Group, University of Delaware, University of Leeds, and Uppsala University for very helpful comments on drafts of this paper.

- Broome, J.: 2001, Are intentions reasons? and how should we cope with incommensurable values?, in C. Morris and A. Ripstein (eds), *Practical Rationality and Preference: Essays for David Gauthier*, Oxford University Press.
- Broome, J.: 2013, *Rationality Through Reasoning*, Wiley-Blackwell, Oxford.
- Chang, R.: 1997, Introduction, in R. Chang (ed.), *Incommensurability, incomparability, and practical reason*, Harvard Univ Press.
- Chang, R.: 2002, The possibility of parity, *Ethics* **112**(4), 659–688.
- Ferrero, L.: 2010, Decisions, diachronic autonomy, and the division of deliberative labor, *Philosophers' Imprint* **10**(2), 1–23.
- Gert, J.: 2004, Value and parity, *Ethics* **114**, 492–510.
- Gert, J.: 2007, Reply to tenenbaum, *Canadian Journal of Philosophy* **37**(3), 463–476.
- Griffin, J.: 1997, Incommensurability: What's the problem?, in R. Chang (ed.), *Incommensurability, Incomparability and Practical Reason*, Harvard University Press Cambridge, MA, pp. 35–51.
- Holton, R.: 2004, Rational resolve, *The Philosophical Review* **113**(4), 507–535.
- Holton, R.: 2009, *Willing, Wanting, Waiting*, Clarendon Press, Oxford.
- Hurka, T.: 1996, Monism, pluralism, and rational regret, *Ethics* **106**(3), 555–575.
- Kavka, G.: 1983, The toxin puzzle, *Analysis* **43**(1), 33–36.
- Kolodny, N.: 2008, The myth of practical consistency, *European Journal of Philosophy* **16**(3), 366–402.
- Paul, S.: 2011, Review of "willing, wanting, and waiting" by richard holton, *Mind* **120**, 889–892.
- Rabinowicz, W.: 2008, Value relations, *Theoria* **74**, 18–49.
- Raz, J.: 1986, *The morality of freedom*, Oxford University Press.
- Raz, J.: 2005, The myth of instrumental rationality, *Journal of Ethics and Social Philosophy*.

- Rosen, G.: 2001, Nominalism, naturalism, epistemic relativism, *Noûs* **35**, 69–91.
- Setiya, K.: 2007, Cognitivism about instrumental reason, *Ethics* **117**(4), 649–673.
- Setiya, K.: 2008, Practical knowledge, *Ethics* .
- Tenenbaum, S.: 2010, The vice of procrastination, in C. Andreou and M. D. White (eds), *The thief of time: Philosophical essays on procrastination*, Oxford University Press.
- Tenenbaum, S.: in preparation, Reconsidering intentions.
- Tenenbaum, S. and Raffman, D.: 2012, Vague projects and the puzzle of the self-torturer, *Ethics* **123**(1), 86–112.
- Thompson, M.: 2008, *Life and action: Elementary structures of practice and practical thought*, Harvard University Press.
- Ullmann-Margalit, E. and Morgenbesser, S.: 1977, Picking and choosing, *Social Research* **44**(4), 757–785.
- Wallace, R. J.: 2001, Normativity, commitment, and instrumental reason, *Philosophers' Imprint* **1**(3), 1–26.
- White, R.: 2005, Epistemic permissiveness, *Philosophical Perspectives* **19**(1), 445–459.
- Wiggins, D.: 1978, Weakness of will commensurability, and the objects of deliberation and desire, *Proceedings of the Aristotelian Society* **79**, 251–277.