

*Reconsidering Intentions**

SERGIO TENENBAUM
University of Toronto

Abstract

This paper argues that the principles of instrumental rationality apply primarily to extended action through time. Most philosophers assume that rational requirements and principles govern in the first instance ~~to~~ *momentary* mental states, as opposed to governing extended intentional actions directly. In the case of instrumental rationality, the relevant mental states or attitudes would typically be preferences, decisions, or intentions. In fact, even those who recognize the extended nature of our agency still assume that rational requirements apply primarily to mental states at a moment in time. Such views try to do justice to the extended nature of our agency by postulating rational requirements that apply in the first instance to plans, policies, and intentions more generally. The paper focuses on the central case of requirements and reasons governing the reconsideration of intentions and argues that these requirements or reasons are either superfluous or invalid.

I argue that a proper conception of instrumental reasoning that applies directly to actions turn out to have surprising consequences. In fact, this conception allows us to see that policies, projects and the like are best understood as instances of extended actions, and that the instrumental requirements that apply to projects and policies are exactly the same as the instrumental requirements that apply to ordinary extended actions. Finally, I argue that the resulting theory of instrumental rationality is a significant improvement over theories that rely on principles governing intentions.

1. Introduction

Most philosophers assume that rational requirements and principles govern mental states or attitudes, as opposed to governing intentional actions¹ directly. In the case of instrumental rationality, the relevant mental states or attitudes would typically be preferences, decisions, or intentions. In fact, most philosophers often assume that

*I would like to thank the following people for invaluable comments on previous drafts of this paper: Chrisoula Andreou, Hagit Benbaji, Michael Bratman, John Broome, Ruth Chang, Philip Clark, David Enoch, Luca Ferrero, Olav Gjelsvik, Matthias Haase, Ulrike Heuer, David Horst, David Hunter, Thomas Hurka, Douglas Lavin, Erasmus Mayr, Jennifer Nagel, Julia Nefsky, Sarah Paul, Jonathan Payton, G. F. Schueler, Rob Shaver, Holly Smith, Michael Thompson, Jonathan Weisberg, and two anonymous referees for *Noûs*. I would also like to thank audiences at the University of California, Berkeley, Harvard University, University of Oslo, University of Oxford, Stockholm University, The Centre for the Ethics at the University of Toronto, The University of York, York University, Leeds Workshop on Practical and Theoretical Agency, Oxford Area Action Theory Group, 2012 Pacific Division Meeting of the APA, the Rationality Workshop at the Hebrew University of Jerusalem, and the Rutgers Value Theory Discussion Group.

principles of practical rationality apply in the first instance to *momentary* states or attitudes of the agent. For example, these assumptions are clearly endorsed by a conception of instrumental rationality based on the principles of decision theory: on such a conception, each momentary decision of a rational agent is supposed to conform to the axioms of expected utility theory. Of course, some philosophers also propose special diachronic principles of practical rationality; in fact, some of these principles are proposed precisely in order to do justice to the extended nature of our agency. However, even such principles are proposed as principles connecting different momentary states of the agent at different times. For instance, Michael Bratman proposes the following principle:

(D) The following is locally irrational: Intending at t_1 to X at t_2 ; throughout t_1 – t_2 confidently taking one’s relevant grounds adequately to support this very intention; and yet at t_2 newly abandoning this intention to X at t_2 .²

Principle (D) is defined in terms of a required relation between two momentary states at two different points in time (t_1 and t_2).³ Similarly, Edward McClennen proposes a diachronic principle “Dynamic Consistency” that relates choice points across different times in an agent’s plans; again, this is in fact a requirement on momentary *decisions* at different points in time.⁴

I think that these widespread assumptions are false, at least for the case of instrumental rationality. In my view, principles of instrumental rationality apply primarily to *actions that extend through time*. However, in this paper, I will not argue for this thesis in its full generality.⁵ Instead, I will look at rational principles that are proposed exactly by those who recognize the extended nature of our agency and try to understand the rational requirements or reasons that arise from making plans, having projects, adopting policies, and, more generally, forming any kind of future-directed intention. Although such views try to account for the nature of extended agency, they are still committed (implicitly or explicitly) to the view that rational requirements apply primarily to momentary mental states.

I’ll argue that future-directed intentions do not play any essential role in our understanding of the instrumental requirements that apply to our agency over time. Instead, I argue that a proper conception of the principle of instrumental reasoning, a conception that applies not to momentary intentions but to extended actions, explains on its own all instrumental requirements that apply to our policies, projects, and similar attitudes. I argue that such attitudes are structurally identical and subject to the exact same basic instrumental requirements as extended actions; norms and principles governing ~~FDIs~~^{future directed intentions} turn out to be at best superfluous.

Thus this paper partly vindicates the claim that the principles of instrumental rationality apply primarily to extended action and that mental states such as future-directed intentions have no significant role to play in a theory of (instrumental) practical rationality. I focus specifically on norms of rationality relating to (non)reconsideration of future-directed intentions because, as I argue below, if there are no norms or requirements directly governing (non)reconsideration of intentions, it is unlikely that there are any kind of norms or requirements directly governing intentions.⁶

The paper begins by considering whether there are any rational requirements or reasons that apply specifically to future-directed intentions (FDIs). I argue that in many of the cases in which we are tempted to say that FDIs make a difference to the rational requirements or instrumental reasons that apply to the agent, it turns out that the FDI is completely irrelevant to the agent's normative situation. The upshot of the discussion is that the best case for instrumental reasons or requirements governing FDIs are cases involving some types of policies and projects.

However such cases turn out to be special cases of extended action, or so I argue. In the remaining sections of the paper, I examine the rational demands that actions, insofar as they extend through time, make on other actions that are located within the same temporal interval. I argue that a more appropriate conception of temporally extended actions can provide the backdrop for a simple but compelling theory of instrumental rationality, a theory that contains no principles governing the formation or execution of FDIs. In fact, the theory relies on nothing beyond a rather uncontroversial version of the classical principle of instrumental reasoning, but I argue that this principle has some surprising consequences when applied to extended action. The resulting theory can thus be applied to acts that were formerly thought to belong to the general category of intentions, such as projects and policies. The resulting theory of instrumental rationality turns out to be quite different from, much simpler, and significantly better than, theories that rely on additional norms and principles governing FDIs.

2. Intentional Autonomy

Let us distinguish three different functions that a future-directed intention (FDI) might have:

[SETTLE] A future-directed intention can settle an issue to avoid reconsideration costs.

Example: I must decide between investing in bonds or mutual funds. I could spend an unlimited amount of time considering the benefits of one option or the other; however, this is not a very fruitful use of my time. So I form a firm intention to invest in bonds.

[LONG] A future-directed intention can allow us to execute long-term plans that require coordination between actions at different points in time.

Example: if I want to go to Machu Picchu, I need to have tickets ready in time, learn some Spanish, etc. Unless I now form the intention to go to Machu Picchu next summer, I will (typically) have no reason to undertake any of these actions.

[RESOLVE] A future-directed intention can make sure that we do not change our minds in the face of temptation or temporary preference shifts.

Example: If I intend to quit smoking, my intention can see me through those difficult moments when someone offers me a much-craved cigarette.

The rationality of forming and carrying out intentions that serve functions [SETTLE] and [LONG] is quite obvious. If we can count on carrying out these intentions, it will be rational for us to form them. Given that we're beings of limited rationality, it would be prohibitively costly for us to reconsider the wisdom of our intentions at every opportunity. Similarly, in the absence of planning and coordination, our menu of options would be severely limited. Since these benefits are still present when I carry out the intention, insofar as I am rational, I expect I'll carry out these intentions if I form them; it is rational for me to form the intentions if I expect I'll carry them out. So it is easy to see how it would be rational for me to form and carry out these intentions. But this does not show that there are reasons or requirements to act on our past FDIs. After all, whatever reasons we had to form the intentions in question are also reasons to carry them out. This is most obviously in the case of [LONG], but not much less clear in the case of [SETTLE]. For, after all, whatever reasons I had to settle on a certain investment will be good reasons to make the investment at a later date. Let us assume that intentions express, or are formed on the basis of, evaluative or normative judgments, such as "I should invest in bonds".⁷ If I was justified in forming this judgment and I had no reason to update, I am justified in retaining this judgment and thus in carrying out my intention.⁸

The preservation of justification through time is less clear in cases under [RESOLVE], but it is also much harder to justify the rationality of forming and carrying out intentions that serve the function of [RESOLVE]. [RESOLVE], unlike [LONG] and [SETTLE], raises the suspicion that forming intentions in these cases is either irrational, or ineffective, or even impossible. To see this point more clearly, we can distinguish three different, but similar, kinds of predicaments that might befall an agent:

- A. Cases in which I predict a preference shift due to an overall change in my evaluative judgments.
Example: I am confident that my startup company is going to make a lot of money in a few years a time. I now prefer that, when my startup company goes public, I donate half of my money to charity. However, I know that I'll be unwilling to give up half my money when I am wealthy. I form now an intention to donate the money to charity when I become rich.⁹
- B. Cases in which I expect temporary preference reversal (change in evaluative judgment).¹⁰
Example: It is always true that I prefer to floss every day than never to floss. This general preference never changes. However, as it comes close to flossing time, I find myself preferring not to floss now over flossing now. Since I know that the result of following these immediate preferences would result in my never flossing, I form an intention to floss always (which at any point in time I prefer over never flossing).
- C. Cases in which, even in accordance with my current preferences, I will prefer not to carry out my intention when it comes time to carry out the intention.
Example: Forming an intention to drink the toxin in Gregory Kavka's renowned toxin puzzle case.¹¹

I'll assume that in all these cases there are no reasons for action that do not flow from my preferences (except, of course, those that the formation of intention generate), but nothing in an argument depends on this. Now, theorists disagree about which of these cases are to be handled by norms governing reconsideration. Richard Holton, for instance, thinks that such norms extend to all cases under (B) and possibly some cases that fall under (C).¹² He is also sympathetic to Edward McClennen's account of resolute choice, and McClennen's account, if taken to validate extending the norms generated by the cases under [LONG] and [SETTLE], would certainly apply to at least some cases like (A).¹³ McClennen himself, like David Gauthier, would extend stability requirements on intentions to (C).¹⁴ Bratman, on the other hand, thinks that one can build an argument that will make similar norms apply to cases like (B), but certainly not to (A) or (C); sticking to one's intentions in such cases would be what Bratman describes as "plan worship".¹⁵

But there are problems with extending any norms arising from cases that fall under [LONG] and [SETTLE] to any of these types of cases. Let us start with (A). [LONG] and [SETTLE] are supposed to be cases that display the importance of FDI in ensuring that agents can continue to carry out projects that they deem important. However, cases under (A) are cases in which I consider a certain project to be important now, but I will not consider it important when it comes time to carry out my intention.¹⁶ It is not clear that the fact that coordination and planning are necessary to achieve an end that I then consider worthless should have any normative pull on me, at least as far as we are concerned only with instrumental rationality. Similar problems apply to (C). (B) seems essentially different since it involves a preference that I will continue to have through time; namely the preference to always floss over never flossing. However, it seems that my choice situation now is essentially different from my choice situation when I formed the intention to floss always, since now I do have the preference not to floss today. It is important to avoid a possible misunderstanding here. Obviously, when I formed the intention I could have anticipated that my preferences would shift in this manner. So it is true even later, when it comes time to floss, that there are no unanticipated circumstances relevant to reconsidering my intention. I did indeed anticipate that my evaluation of the various courses of action would change. But under the assumption that, in general, my current preferences or evaluations are relevant for what I am instrumentally required to do now, I should have anticipated that I would at a later time be required to reconsider the intention; the fact that my intention was formed on the basis of evaluations or preferences that I now reject seems to be a paradigmatic case of a circumstance in which reconsideration is called for.

None of these arguments are decisive. One could claim that the kind of coordination and planning involved in the standard cases also applies to coordination and planning in light of the fact that human beings constantly change their preferences, especially in the kind of cases in which we seem to be prone to hyperbolic discounting.¹⁷ More generally, one might argue that [RESOLVE] cases are not importantly different from [LONG] and [SETTLE] cases, and whatever reasons or requirements apply to the latter cases would also apply to the former.

In sum here are two theses that one might defend:

Intentional Autonomy: Forming FDIs generates, directly or indirectly, new reasons, or rational requirements (or rational principles) that do not reduce to non-intention based reasons or requirements or to the reasons generated by the consequences of forming the FDI.

The Extension View: The new reasons or requirements also apply to [RESOLVE] and explain the rationality of sticking to our resolutions in these cases.

The intentional autonomy thesis implies that intention has its own “rational life”. In forming an FDI, we are bound by certain normative standards that are internal to the state. This should be distinguished from the (trivial) claim that forming an FDI will lead us to act in certain ways, and the outcomes of such actions will generate new reasons. I intend to go to Paris next month, and thus I buy tickets to the major attractions, make non-refundable hotel reservations in Paris, etc. I now have reasons to go to Paris that I didn’t have before. But these reasons are generated by the intention extrinsically; they are what Bratman calls “snowball effects”¹⁸ of my intentions. No one doubts that future-directed intentions could change our reasons by changing the facts that can be reasons for our actions.¹⁹

The *Extension View* is just a stronger version of *Intentional Autonomy*. I will argue that both views are false. Obviously, if *Intentional Autonomy* is false, so is the *Extension View*. But it is worth presenting the two views separately because the *Extension View* faces some particular difficulties on its own. I’ll present theories that accept the intentional autonomy thesis as two-tier models of rationality. On some theories, the existence of two tiers of evaluation might turn out to be superfluous, but it will be easier to present my arguments by generalizing this form. In this way, we can include under “two-tier theories” virtually any theory that accepts that there are independent rational requirements on the (non)reconsideration of intentions or independent reasons not to reconsider intentions. Such theories evaluate the agent’s rationality in reconsidering a relevant intention (and the subsequent choice) using the following general structure:

The Two Tier Model (TTM)

First Tier: Evaluates the rational status of the agent’s habits, dispositions, principles, etc. of (non)reconsideration.

Second Tier: Evaluate the choice of these actions by whether they conform to principles, habits, dispositions, etc. that were judged to be rational by the first tier.²⁰

In a nutshell, the first tier imposes general requirements on the formations of principles and dispositions of intention (non)reconsideration and the second tier evaluates conformity with these principles and dispositions. As we said above, it is common ground that non-reconsideration is sometimes rational; for instance, since it is rational at times not to reconsider one’s evaluative or normative judgments, it is also rational not to reconsider the intentions that are based upon or express these judgments. However, given the coordinating functions of intentions,

one might think that there are reasons internal to the activity of intention formation that provide us with special criteria of when it is rational to revise an intention. When asking ourselves whether we need to revise our judgment, we are concerned only with whether we have reason to think that it is defective, and, perhaps, with whether our resources are better spent elsewhere. On the other hand, were I to reconsider my decision to invest in stocks, I would be undermining the very point of having formed the intention in first place. And were I to constantly reconsider my intentions for long-term projects, it would make it unfeasible for me to engage in them. So it seems plausible that the norms of (non)reconsideration and revision for intention are independent of the norms of (non)reconsideration and revision that apply for judgments about how we should act; these norms of (non)reconsideration are not mere consequences of these judgments. So focusing on requirements governing reconsideration is not picking on a random requirement in the intentional autonomist treasure chest; requirements governing intention reconsideration are likely to be the crown jewels of the advocate of intentional autonomy.²¹

3. Reconsidering Intentional Autonomy; Rejecting the Extension View

How do planning and coordination give rise to rational norms of reconsideration? The function of intentions we found in [LONG] and [SETTLE] seem to result in the following considerations in favour of special norms governing intention reconsideration.²²

[STABILITY] We need stable intentions so that (i) we can engage in long-term projects and so that (ii) we are predictable and can coordinate with ourselves and others. Thus we ought not to reconsider future-directed intentions at every opportunity

and

[DECISION] We need to be able to decide without constant deliberation. Thus we ought not to reconsider FDIs at every opportunity.

These seem quite intuitive reasons to accept that it would be rational to have limits on reconsideration. But what exactly are these limits? Let us look first at what sorts of demands [DECISION] makes on us. Of course sometimes in exigent circumstances deliberation is certainly a bad idea. If I know that my banker will come at 9:00am to write down my decisions about how to invest the money I have inherited, and she'll leave at 9:01am and invest my money however I have instructed her (or she'll simply put it under her mattress if I have not instructed her to invest it in any particular way), I'd better do most of my deliberation before 9:00am. And once I come up with a decision, I'd better stick to it during the minute-long meeting, so that my money does not end up at the bottom of my banker's mattress. If I decided that I should put all my money in index funds, reconsidering this intention at 9:00am, insofar as it involves deliberation, would obviously be a bad idea. Sticking to my intention to put all my money in index funds seems to be a better action than deliberation at 9:00am.

But why is it better? Presumably because I have good grounds to think that my previous decision is likely to have been a good one; at least, I expect it to be better than a random guess. Suppose we have a similar setup except that I am not aware that I'll only have one minute to make up my mind; I actually think that my banker will sit with me and talk to me about all my options, give me further opportunity to think about my options, and be there with me for as long as I want. I spend the day before thinking about the options and, so far, I think the best option is to go for bonds. However, I do not form any intentions; I decide it is better to "sleep on it", and share my views with the banker. I arrive at the bank at 9:00am without having put any more thought into the issue, and as I step into the bank, my banker suddenly announces that I will have only one minute to make my decision. I have here just as much reason to follow my earlier assessment as I had not to reconsider my intention in the case where I did form an intention.²³

On the other hand, suppose that I form an intention, and there are no exigent circumstances which prevent me from deliberating, or which make deliberation too costly. Suppose I am facing the same issue of how to invest the money I inherited. The day before the meeting, I form the intention of buying into an index fund. Now I am waiting for my banker, and I expect that I'll have at least half an hour before she can see me. Would it be irrational for me to reopen the issue? There's really no new information, no unexpected circumstances (suppose I know that I generally have to wait about half an hour before I meet my banker), but I am bored, I don't know what to do, and reconsidering my intention seems to be a perfectly good way to spend this half hour.²⁴ It couldn't possibly be a failure of rationality to reconsider my intention in these circumstances.

Of course, it would be a horrible fate to spend most of one's life reconsidering one's decisions; it would be absurd if a theory of rationality required us to use all our leisure time, let alone all our time, in the service of better deliberation. Since the completion of many of our ends would be hampered if we were to deliberate too much, we are all already under a requirement not to reconsider too much. Given that I have among my ends exercising, spending time with the children, watching cartoons, etc., I cannot coherently also engage, at the same time, in endless deliberation. A theory of rationality must allow that, in certain circumstances, we are *permitted* not to reconsider an intention even if we expect that our decision might improve and nothing of much value will be lost if we do. However, the same kind of consideration cannot ground a *requirement* not to reconsider intentions in non-exigent circumstances; after all, reconsidering in any such particular occasion will not prevent me from having enough time left for a quite a bit of exercising, spending time with the children, and watching cartoons.

But we should also note that similar permissive norms govern any attitude that it makes sense to "reconsider". Take belief, for instance. It would also be a sad kind of life spent reconsidering our beliefs all the time. There is thus no rational requirement to reconsider our beliefs at every opportunity, or to always look for further evidence; we are often permitted to settle on a belief. But there is also no general requirement not to reconsider a belief²⁵ unless there is enough change in information, etc.²⁶ In fact, one can hardly imagine that an attitude that one can

form in response to reasons would not entail a general (defeasible) permission not to reconsider whether the attitude is warranted. After all, a rational agent would, in reconsidering, suspend judgment on whether there are reasons to form the attitude and thus no longer have the attitude in question. Such rational attitudes would be no more than momentary, self-extinguishing ephemera. If, in forming a belief that **p**, I were *required* to immediately reconsider whether there were good reasons to believe that **p**, I would have to suspend judgment on whether there were good reasons to believe that **p**, and thus, insofar as I am rational, withhold judgment on **p**. I would thus give up my belief as soon as I formed it.

Bratman claims that intention is a kind of mental state that involves a disposition to “resist reconsideration”. A natural way of understanding this claim is to take “disposition” here to refer to a normative disposition; our “default position” should be not to reconsider an intention we have.²⁷ If I am correct then this claim is misleading; at most we can say that intentions typically don’t *demand* reconsideration.²⁸ Cases that fall under [SETTLE] can ground *permissions* not to reconsider FDIs in certain situations; yet, even this turns out to be superfluous. It is in the nature of forming rational attitudes that there is no general requirement to reconsider whether my reasons for forming the attitude were good ones.

However, it is true that were we not to exercise enough of these permissions, constant deliberation and revision would undermine our capacity to pursue any long-term ends.²⁹ So even if (DECISION) fails to establish a requirement, (STABILITY) might succeed. There must be a more general requirement that prevents us from becoming “reconsideration monsters”. Reconsideration is an activity that interferes with many of our goals; we need to ensure that we don’t reconsider so much as to prevent our engagement in other important activities. But, as we said above, the pursuit of these goals already imposes limits on reconsideration; in particular, a prohibition that extends as far as necessary for their successful pursuit. The instrumental requirements that follow from these various goals would seem to provide enough stability to pursue long-term ends with sufficient consistency. On no one’s view are plans never to be cancelled, and not even the most rational being can always be relied on to proceed as expected. Moreover “snowball effects” also conspire to ensure that a rational agent will be reasonably stable; after I buy a non-refundable ticket to Paris, it becomes extremely costly to change my mind and spend my weekend in Lubbock instead, even if ordinarily I am indifferent between vacationing in Paris or Lubbock.

Finally, it is worth noting that we do not need a rational requirement to ensure the ends of coordination and stable pursuit of the long-term ends. What we need is that agents can be *expected* to proceed as they intended (often enough), not that they be *required* to proceed as they intended (in the right circumstances). But if at t_0 I intend to ϕ at t_1 and nothing interferes between t_0 and t_1 , then I will still be intending to be ϕ -ing at t_1 , which means that, in the absence of interferences, I will be ϕ -ing at t_1 . Of course, reconsidering the intention is a possible source of interference, but the fact that I am not required to refrain from reconsidering does not entail that I expect that I will reconsider (let alone that I will reconsider and change my mind). Thus the mere absence of a requirement not to reconsider (or

something similar) does not prevent rational agents from coordinating and forming stable plans. Snowball effects, our other ends, the other demands on our time and mental resources, and even the very fact that in intending to ϕ I (typically) acquire a belief that I will ϕ , already conspire to make our intentions stable.

Nonetheless reconsideration has a problematic structure. Quite often reconsidering is not too costly, and for (almost) any particular plan, abandoning just this plan will not undermine more general ends. But reconsidering in too many cases can have a devastating cumulative effect: were I to reconsider my intentions at every permitted opportunity, I would forgo pursuing many of the ends I care about. And were reconsideration to lead me to revise plans often enough, my life would be a pathetic alternation of momentary or soon-to-be-abandoned pursuits. On the one hand, no *particular* intention must “resist reconsideration”; the requirement not to reconsider too much applies only to the total *set* of one’s intentions. On the other hand, we can only satisfy this requirement if we avoid reconsidering particular intentions, none of which we are required not to reconsider.

TTM might be thus defended as determining what counts as rational agency for agents who face this kind of problematic structure. Given that on any particular occasion there is no reason to refrain from reconsidering, an agent will risk reconsidering too much, and thus jeopardize her plan and her ends, if her habits and dispositions are not such that, sufficiently often, her intentions are reconsideration-resistant, or, in general, if there are no reasons or requirements pushing against reconsideration. This problematic structure is widespread; we can give endless examples of activities such that, at least apparently, at various particular instances it would be rational to engage in such activity, but also such that it would be disastrous to engage in this activity in every such instance. For instance, as I look at my twitter feed, I could just look at one more tweet before preparing my class. For each tweet, it seems true that reading just this one will not jeopardize my class preparation.³⁰

One might say that given the general nature of the problem, we should expect that a solution to it would not generate any requirements that are *specific* to intentions. But this is obviously wrongheaded; if anything, the generality of the problem speaks in favour of TTM. After all, all such cases seem to be instances of the derailment of long-term plans, projects, etc. by conflicting momentary (or short-term) ends and preferences. In other words, the problem seems to be precisely a problem about how general intentions should guide our momentary actions, and TTM is a proposal about, *inter alia*, how a (limited) rational agent implements general intentions. Furthermore, these are cases in which rational agency seems to require stable intentions. An ideally rational agent facing such choice situations would not be prevented from implementing her policies (and achieving various long-term ends) by the structure of local choices. And it seems that she can only implement her policies by somehow avoiding reconsidering her policies and long-term ends too often. In sum, even if all my arguments so far are successful, rational agency in the face of this kind of problematic structure might involve being responsive to reasons and requirements governing intention (non)reconsideration. So my next step is to argue that TTM (and thus any account that relies on such reasons or

requirements) is neither the only nor the best account of rational agency in such choice situations.

Before we move on, it is worth noting that the Extension View is not faring as well. What we need in [RESOLVE] is a *requirement* to stick by our intentions, not just a permission. The various reasons to allow for non-reconsideration seem unable to produce the requirements that the extension view demands; the extension view needs to *prohibit* certain kinds of reconsideration, not just permit non-reconsideration. Moreover to the extent that there are rational requirements that prohibit too much reconsideration, these rational requirements are conditional on the end they serve. That is, if the need for coordination generates a rational requirement to limit how much one revises one's intentions, it limits only to the extent that such revision would be incompatible with the general aim of allowing for interpersonal or intrapersonal coordination. But succumbing to temptation, especially if I do this in a predictable manner, will not prevent me from coordinating with myself or others. Of course, I might have other reasons not to succumb to temptation on this particular occasion, and, as we'll consider in more detail in section five, I have long-term ends that I fail to pursue efficiently if I succumb to temptation *too often*. But coordination is not fundamentally impeded if I break my resolutions sometimes (especially if it happens in predictable ways). Similarly, my actions will be stable *enough* even if, from time to time, I succumb to temptation.

More importantly, a similar problem haunts any pragmatic defense of the extension view, any defense that appeals to the overall benefits of sticking to one's resolution in the life of a rational agent. For in all such cases it will be similarly difficult to move from the realization that sticking to resolutions *often* or *often enough* is beneficial to the claim that it must be done in this particular occasion. Just as in the case of coordination, we will at most be in position to generate a permission to stick to one's resolution and possibly a requirement not to break resolutions *too often*. But it is odd, to say the least, to form a resolution not to drink tonight, while accepting a norm that I should stick to my resolutions *sometimes* or even *often*; it seems that if I do this, my sober colleagues would be right to say that I didn't really *resolve* not to drink tonight; at best, I came to the conclusion that it might be a good idea not to drink tonight, or that not drinking is something I should do from time to time. In the language I will introduce shortly, resolutions tend to be *strict*, but strict policies or rules are rationally adopted for instrumental reasons only in very special circumstances. Unless one can show that it is psychologically (nearly) impossible to follow through less strict policies in such cases, instrumental reasons cannot warrant a strict resolution.³¹ The TTM model has the best chance to succeed in cases in which there is no similar strict requirement to always follow your general intention; the TTM model is best suited to account for the rational agency in relation to what I will call *vague* policies.

It is important to recognize some limitations of this argument. In arguing against the extension view, I have not shown that no intention-specific requirements apply to [RESOLVE] cases; I've only tried to show that we cannot *extend* certain justifications for intention-specific norms to these cases in the appropriate manner. It is entirely possible that we can argue for intention-specific requirements that apply

to [RESOLVE] on completely different grounds. Unfortunately, I cannot here delve into every possible argument of this kind.³²

4. Extended Action and Policies

Let us go back to our question of how a rational agent chooses in situations with the problematic structure discussed in the previous section. Let us first distinguish between two kinds of intentional actions: “gappy” and “gapless” ones. Actions that stretch through time will typically have many parts that are (shorter) actions performed as a means to the (longer) action in question. My buying some milk has as its parts walking to the store, getting the milk out of the store fridge, taking money out of my wallet, etc. In some cases, during the interval of time in which I am performing an action, all the actions I perform throughout its duration are means to the larger action; these are the gapless actions. Swimming from A to B might be like this; while I am swimming, everything I do is a means to swimming from A to B.³³ But the same is not true of taking a walk; taking a walk is gappy. Suppose I go for a walk, and as I am engaged in this activity, I stop and look at a beautiful suit on display in the store. I am still taking a walk when this happens, even if stopping to see the display is not a means, properly speaking, to my taking a walk.³⁴ Once we introduce the notion that actions are gappy, there is no principled reason to limit the width the gaps between the proper means to the actions. There are very few, short gaps in cooking an omelet, larger gaps when taking a walk, and even larger gaps in writing a book. But once we note a continuum of indefinite length of gaps between proper means, various things that do not spring to our mind as examples of actions seem to naturally belong in this continuum. So we might think that the peripatetic philosopher is engaged in an action much like taking a walk; it is in fact the action of taking a walk (and contemplating) with much wider gaps and stretched through a much longer period of time.

The structure of activity of the peripatetic philosopher is not significantly different from the structure of the activity of someone who has adopted a policy. In both cases, there is (we suppose) a reason to engage in certain characteristic activities (walking and philosophizing, for the peripatetic philosopher; running, swimming, etc., for the person who has a policy of exercising), and the agent acts on this same reason at various occasions. Success in being a peripatetic philosopher involves engaging in these activities often enough, and similarly success in implementing one’s policy of exercising involves engaging in its characteristic activities often enough.

My suggestion is that any kind of policy, project, long-term action, etc. can be understood in these terms; that is, as a continuous (though “gappy”) action. However, I am not concerned here about the metaphysics of actions or policies; all I want to argue for is that from the point of view of the theory of rationality, there are no differences between actions and policies.³⁵ Perhaps actions require certain bodily movements, while policies don’t. Perhaps actions are events or processes while policies are states. I do not want to take a stand on these issues. My sole contention is that such differences, if they indeed exist, do not result in distinctions that are relevant for a theory of practical rationality. In particular, a long-term

action will consist in part of the undertaking of many momentary, or at least short-term actions. The rationality of the set of momentary actions will be in part (and perhaps fully) a function of how the momentary actions contribute to the execution of the long-term action and its ends. Pre-heating the oven is rational (partly) because it contributes to my baking the cake. Similarly the “gaps” in the production of the cake will be judged to be rational in part due to their relation to the larger action; whatever I do while I am baking a cake should not interfere with the production of an acceptable cake.³⁶

For ideally rational agents engaged in a long-term action ϕ , some momentary actions in its interval are constitutive or instrumental means (where “means” is conceived broadly) to ϕ , while the gaps between these momentary actions will include only actions that are not incompatible with (or don’t seriously endanger) the completion or continuation of ϕ . If we then treat a policy as an ordinary action, acts in accordance with one’s policy will be similarly constitutive means to it, and actions in the gap between two instances of execution of a policy must be such as not to be incompatible with (or not seriously endanger) the maintenance of the policy. In the case of ordinary actions, the rationality of momentary actions and gaps is evaluated in relation to the *completion* or *maintenance* of the larger action. This pair of notions is necessary to mark the different relations of two different types of action that roughly correspond to the referents of what Zeno Vendler calls “activity terms” and “accomplishment terms”.³⁷ Some actions (activities) are completed at any point at which they are interrupted (if I was running, I have (gone for a) run); some actions (accomplishments) are only completed at the end of the process (I can’t infer that I have built a house from the fact that I was building a house; I have only built a house if the process of building a house was completed).³⁸ Policies, on the other hand, are always activities, and thus the rationality of a policy must be understood in terms of its maintenance or perseverance. Of course, we should not conclude, at least not by definitional fiat or by the examination of the grammar of certain sentences, that instrumental rationality requires that once we engage in an activity, we not abandon it. What instrumental rationality requires is that we pursue the means to our ends. In the case of activities, the constitutive and instrumental means for the end of the activity are for its continuation or perseverance, not for its completion. The instrumental requirement only applies as long as you still have the end, while you are still engaged in the activity.

One might worry that, unless we postulate a diachronic requirement that I continue an activity after I start it, instrumental requirements cannot apply to activities; after all, if I do not pursue necessary means for my ϕ -ing, then I am not ϕ -ing, and thus failing to pursue the means to what I am doing. It is true that there could not be any non-diachronic instrumental requirements for activities if we assumed that non-diachronic requirements must be momentary. In formal epistemology, a view that defends the existence of diachronic norms of rationality (let us call it “the diachronic view”) is contrasted with the “time-slice” view, a view according to which norms of rationality apply fundamentally to instantaneous time-slices of the agent.³⁹ But at least on one understanding of “diachronic norms”, this contrast need not be exhaustive. A diachronic norm N of X-ing⁴⁰ can be defined

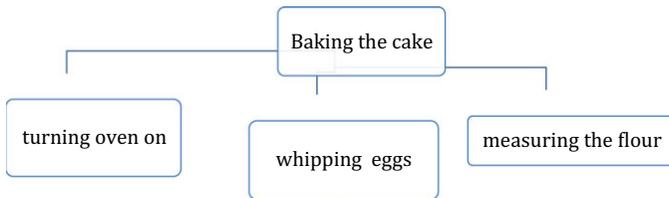
as a norm such that whether an agent complies with N at t depends on facts that obtain at some time t_1 such that $t_1 \neq t$.⁴¹ But the time-slice view follows from the denial of the possibility of diachronic norms only if “ t ” refers to moments; if the formulation is neutral between moments and intervals, a “synchronic” norm is not necessarily a time-slice norm. Given that our actions and pursuits extend through time, when “X-ing” refers to doings, or things like “pursuing an end”, a norm that is synchronic to X’s occurrence could depend on facts that are not circumscribed within any particular moment in time without depending on facts that do not overlap the time of X’s occurrence. Of course, if one assumes that the rationality of an agent at an interval must supervene on the rationality of the agent at each moment within the interval, then rejecting the diachronic view implies accepting the time-slice view. But I will argue below that this assumption is false.⁴²

At any rate, one can be *pursuing* the end of ϕ -ing even while *at the same time* failing to take the necessary means to ϕ -ing, as long as pursuing an end extends through time. Let us look at two distinct cases of such a failure. The first case is an adaptation of Anscombe’s understanding of the “contradiction” of the expression of an intention in action. If I am replenishing the water supply and at the same time I am making holes in the pipe with an axe then I am being instrumentally irrational in pursuing the end of replenishing the water supply.⁴³ Although it seems a rather bizarre agent who would act irrationally in this way, we can think of less far-fetched examples. Suppose I am trying to run a six-minute mile. A six-minute mile is an accomplishment and I can be instrumentally irrational in ordinary ways in bringing this about; I can give in to temptation and slow down at various points, and thereby end up taking longer to finish the mile. But a similar “syndrome” could be present if I am running at a six-minute mile pace. Here too at various points I could give in to tiredness or distraction and slow down my pace. And this could all be happening while I am simultaneously acting in ways that are intelligible only if taken as means to my (failing) pursuit of running a six-minute mile; I could be at the same time putting a great deal of effort in breathing in the way necessary to keep a six-minute mile pace (but superfluous if I am trying to run at slower paces). At those moments I would be instrumentally irrational exactly because I am pursuing the end of running a six-minute mile pace yet not taking the necessary means to it.

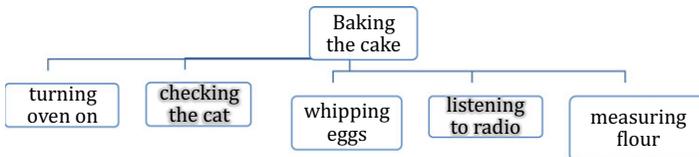
But the possibility of instrumental irrationality for activities is even more apparent in some gappy activities. For instance, suppose I am keeping a rope tied to a pole at a certain level (we need it to be, say, always roughly within 10cm of one meter above the ground for our jumping competition). The rope slowly slides down the pole so I need to keep pushing it up. But because it can be a bit further up or down, I can push it up and then do something else quickly before I need to push it up again. So I could check the beautiful cat across the street before I need to give the rope another tug. But here I could linger too much, and irrationally fail to take the necessary means of my pursuing the end of keeping the rope at around 1 meter. Yet there is no question that I am still engaged in the activity; there is no other (explanatory) reason why I am holding on to the rope other than

the fact that I am trying to keep the rope tied to the pole in this way. The gappier the activity, the easier it'll be to make room for its irrational pursuit. In keeping a diary, I can keep postponing writing my next entry, until it is no longer true that I am keeping a diary. Just as one can be instrumentally irrational in relation to the action of writing a book, one could be instrumentally irrational in relation to the end of keeping a diary. I skip a day or two, then more, etc. until at some point I am not succeeding in the pursuit of my end of keeping a diary. And as I write the last entries after long intervals, I am still pursuing the end of keeping a diary while not taking adequate means to it.

We can now propose an alternative to TTM's treatment of policies and projects that exhibit the problematic structure discussed in the previous section. I'll call this alternative the "Policy as Action Model" (PAM). Let us start with examining how instrumental requirements apply to a "mundane" long-term gappy action such as baking a cake. We then see how to extend it to a policy such as "exercising regularly" on the assumption that policies should be understood as just instances of long-term gappy actions. We can represent a long-term action, such as baking a cake, by means of the following diagram:



The chart only shows actions that are instrumental to baking a cake. But baking a cake is gappy; not all my actions while I am baking are cake are means to baking the cake. I might stop to listen to some piece of news in the radio, look out and check whether the neighbour's cat is destroying my plants, etc. The following is a more detailed representation of a gappy action:



Obviously, there is a difference between the rational relation of the larger action to its means, and the relation of the larger action to its other parts. After all, my reason for measuring the flour is that I am baking a cake, but my reason for checking the cat is certainly not that I am baking a cake; or, in other words, I measure the flour for the sake of baking the cake, but I do not check the cat for the sake of baking the cake. But we can say that the smaller actions are under the rational control of the larger action given that, insofar as I am rational, the amount of time and effort that I spend on these other activities is constrained by my interest

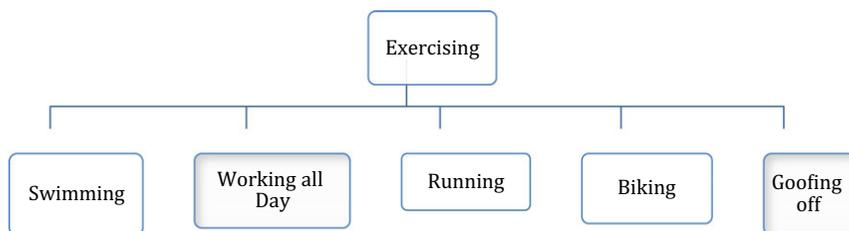
in performing the larger action. I should only listen to the radio as long as it does not interfere with my baking the cake.

Most important for my purposes is the following contention:

(SUFFICIENCY) For my actions to be instrumentally rational in relation to the end of ϕ -ing (baking the cake), it is sufficient that I ϕ -ed (baked the cake) through my actions in expected ways.⁴⁴

Of course, “that I ϕ -ed through my actions in expected ways” cannot express a *necessary* condition for my rationality with respect to ϕ -ing; after all, stepmotherly nature might have conspired so that my best efforts were insufficient for my having ϕ -ed. But one might complain that (SUFFICIENCY) fails to capture a sufficient condition for instrumental rationality. For if I did bake a cake, but I did so in a way that, as I could have predicted, set the house on fire (if, for instance, I prepared the caramel filling by burning sugar laid on a wooden table with a flamethrower), I would have not been rational; a cake is not worth the destruction of my abode. But (SUFFICIENCY) just says that I was instrumentally rational *in relation to the end of baking a cake*. This does not preclude my actions from being instrumentally irrational because they were incompatible with other ends (such as my end of keeping my house in good living conditions).

If having a policy is an ordinary action, then the same forms of rational control hold between the policy and its instances as between the larger actions and its parts. Let us take, for instance, a policy of exercising regularly. PAM posits a similar structure of direct and indirect control of the policy over actions as in the following:



Assuming that (SUFFICIENCY) is true, we get an important consequence of PAM for the theory of instrumental rationality. The claim that policies exhibit the same rational structure as actions implies that a suitably modified version of (SUFFICIENCY) is true of policies. In particular PAM is committed to the following claim:

(SUFFICIENCY_P) For my actions to be instrumentally rational in relation to the end of executing my policy P (e.g. exercising), it is sufficient that I execute the policy (e.g. exercise enough) through my actions in expected ways.

Here is a place in which the intentional autonomy thesis might diverge from PAM. PAM sets up a sufficient condition of instrumental rationality in relation

to a plan that is independent of any requirement that is specific to intentions; (SUFFICIENCY_P) is just the application of the principle of instrumental reasoning (or, more specifically (SUFFICIENCY)) to policies. A two-tier model seems to play no role in understanding whether an agent satisfies (SUFFICIENCY_P) in relation to a policy.

The full extent of possible disagreements between PAM and the two tier model can only be assessed if we also formulate necessary conditions for the rationality of a policy according to PAM, and compare the detailed theory with various versions of the two tier model. So far the only constraint we have is that these necessary conditions are the same as those governing the instrumental rationality of actions. We could say, roughly, that an agent is rational in pursuing the end of ϕ -ing only if she can reasonably expect that her actions result in her having ϕ -ed or in her continued ϕ -ing. I'll not undertake this detailed comparison between PAM and TTM. Instead, I will look at some very general versions of the two tier model and look for some possible sources of conflict with PAM. I'll argue that this partial comparison is sufficient to show the superiority of PAM.

Let us look at cases in which a certain policy can potentially apply. For instance, the policy "Drinking moderately", or a general policy against drinking that allows for occasional exceptions, puts no constraints on my action if I am, say, walking on a campus in which no alcohol can be found. It'll apply if I am at a party at which alcohol is available in large quantities. Once a policy applies, TTM assesses the rationality of (non)reconsideration at two levels.⁴⁵ Instances of (non)reconsideration are deemed rational if they are the outcome of reasonable principles, habits, or dispositions of (non)reconsideration.⁴⁶ Dispositions are deemed reasonable if they reliably produce the right kind of actions (for instance, actions that tend to bring it about that our ends or preferences are satisfied). Let us now distinguish two versions of the two tier model depending on how they read the first tier:

(TTM Strong): Instances of (non)reconsideration are deemed rational if, and only if, they are the outcome of reasonable dispositions of (non)reconsideration.

(TTM Weak): Instances of (non)reconsideration are deemed rational if they are the outcome of reasonable habits or dispositions of (non)reconsideration.⁴⁷

TTM Strong only evaluates an agent as rational if she is manifesting a reasonable disposition. An agent might fail, according to TTM Strong, in two ways. On the one hand, some of her dispositions of (non)reconsideration might fail to be reasonable, in which case *all* manifestations of such disposition would be irrational. On the other hand, she might have reasonable dispositions that she fails to manifest in certain circumstances. She might act "out of character" in some situations and (fail to) reconsider when, were her actions to manifest her reasonable dispositions, she would not (would) have reconsidered. TTM Weak, on the other hand, deems all manifestations of a reasonable disposition rational, but is silent on those cases in which the agent's disposition is not reasonable, or those in which her reasonable disposition is somehow inert.⁴⁸

5. PAM vs TTM Strong

Let us distinguish between two kinds of policy: loose and strict.⁴⁹ Strict policies are those that they call for the performance of the relevant action in every occasion in which the policy applies; a loose policy allows for occasional exceptions. So, for instance, for many people a policy of loyalty to their partner is a strict policy. On the other hand, one can have a more Leonard Cohen-like fidelity policy, which only requires you to be “faithful, Ah give or take a night or two”.⁵⁰ Using Bratman’s distinction between rejecting a policy and blocking a policy, one can say that strict policies can be rejected, but not blocked.⁵¹ Larry may decide that he is no longer interested in being loyal to his spouse and no longer adopt a policy of loyalty, but insofar as he maintains a strict policy, he cannot rationally continue having the policy while deciding that it does not apply on Tuesdays, or in Paris, etc. The Leonard Cohen version of a fidelity policy also illustrates that loose policies can be of two kinds. The phrase “give or take a night or two” suggests that this version of the policy does not specify precise exceptions. Another loose policy might be more precise in specifying when certain exceptions are allowed. So, for instance, one can have a policy of not drinking more than two pints of beer per evening, or a policy of drinking moderately, or just a general policy not to drink that allows for occasional exceptions. The first is loose, but precise, policy, the second and third one are loose and vague.⁵²

Now TTM Strong and PAM cannot give different verdicts to strict policies, for each must classify every action that fails to conform to such a general policy as instrumentally irrational. The same holds for loose but precise policies, as any violation not covered by a precise exception will be deemed instrumentally irrational by both views given that such violations are straightforward violations of the principle of instrumental reasoning. If I have a strict policy never to talk to strangers, any time I talk to a stranger will count as a violation of the policy. However, we certainly need not appeal to special principles generated by TTM to explain this irrationality. If “never talking to strangers” is one of my ends and I talk to a stranger on my walk, I have failed to pursue the necessary (constitutive) means to my end; this is a trivial instance of violating the principle of instrumental reasoning.

It is in the case of vague policies that the potential for disagreement comes up. Suppose I have a policy of having fruit for dessert instead of chocolate. Suppose my reason for adopting this policy is health-related: I like chocolate better, but I am concerned about my health. The best policy in this case would then be a vague one; this is a policy whose application a rational being would block in certain circumstances. There are various ways of representing this policy, but for our purposes, it will be the best to think in the following terms: the policy in question is just a general policy “Eat fruits for dessert”. Since this is not a strict policy, I allow myself some exceptions; sometimes I can block the policy without thereby rejecting it. In other words, I can decide not to follow the policy today without in any way giving up on, or even weakening, my commitment to not eating dessert. I’ll assume that the policy is executed if nothing interferes with its execution. In our

example, in the absence of interference, when it comes time for dessert, I'll eat fruits if available, or nothing at all.⁵³ A typical case of interference, and for our purposes the only relevant one, is reconsideration. In fact, given that presumably enjoying sweets from time to time is one of my ends, I *should* reconsider from time to time; the only open question is whether there are some occasions on which reconsidering would manifest irrationality or some kind of rational failure. Finally, I'll assume that whenever I reconsider I choose not to follow the policy; this is, of course, a simplifying assumption and nothing in the argument depends on it.

Notice that this policy is not very different from a policy such as "eating fruits and vegetables regularly". Here too TTM might be thought to provide a plausible understanding of the rational implementation of such a policy; a rational agent will have dispositions that guide her in implementing such a policy, and she is rational if her actions, in the appropriate circumstances, are manifestations of such dispositions. We determine whether deviations from the dispositions are rational by first determining whether reconsideration of the course of action recommended by one's general disposition is rational. Suppose the relevant disposition would have led me to eat fruit in the morning, but I decide not to do it. We could now represent this deviation as a case of reconsidering whether I should implement the policy in this way, and evaluate the rationality of potential deviations as the rationality of reconsidering (and deciding on the basis of such reconsideration). The same could be said of any vague policy, such as exercising regularly, driving carefully, treating students politely and so forth. Since PAM and TTM would give divergent verdicts of rationality mostly with respect to such deviations,⁵⁴ the discussion of this particular case should generalize to the implementation of *any possible vague policy*. Since we already established in the earlier sections that the special reasons or requirements proposed by TTM could apply *at most* to the implementation of policies, once we show that PAM fares better in explaining the rational principles governing the implementation of such policies, we have established the general superiority of PAM, and therefore shown that the purported reasons or requirements introduced by TTM are all either superfluous or incorrect.

Let us return to our policy of eating fruits for dessert. According to PAM, as long as my actions exhibit the overall pattern in which I don't eat too much chocolate, there is no further question of the rationality of my conduct. (SUFFICIENCY_p) implies that the existence of a pattern of eating in accordance of the policy guarantees that I was instrumentally rational in relation to the policy.⁵⁵ On TTM Strong, on the other hand, at least in some cases in which I had sweets for dessert, I might have failed to act in accordance with rational habits and dispositions of reconsideration. Suppose, for instance, I have a reasonable habit of only reconsidering my intention not to eat chocolate for dessert in certain unusual circumstances;⁵⁶ when the chocolate is especially good, or goes particularly well with this liqueur, etc. Suppose now that this is not one of these evenings, but I reconsider anyway, and in light of my current preference for eating chocolate, I simply do it. On this version of TTM, this is a case of irrationality, while in PAM, there's nothing irrational in this action *per se*. Doubtless, if this kind of reconsideration keeps happening, the

general *pattern* of activity could turn out to be irrational. But if PAM is correct, this particular action cannot, on its own, ground any assessment of instrumental irrationality in relation to this policy.

But if the agent's "dessert behaviour" still conforms to his vague policy, why should such an action count as irrational? Doesn't (SUFFICIENCY_P) correctly evaluate such patterns of activity as rational relative to the policy? Gauthier gives an argument for something like TTM Strong (in a somewhat different context). In considering a (vague) policy to have fruit for dessert, Gauthier argues that I could not choose chocolate in a particular occasion simply because at the time of the decision, I want (prefer) to have chocolate on this occasion:

In considering my choices between chocolate and a fruit, I am supposing that if it were rational for me to deliberate on some basis for any such choice, it would be rational for me to deliberate on that basis for every such choice. For there is nothing to distinguish one choice from another . . . [Thus] choosing on the basis of one's proximate preferences is not rational.⁵⁷

Now this kind of consideration seems to lend support to TTM Strong. After all, if one doesn't rely on dispositions that can ensure a certain regularity of action in various choice situations, one would have to choose exceptions simply on the merits of the particular situation. But if I am ready to make an exception in a case without relying on any special features of the situation, then, by parity of reasoning, I would be willing to make an exception to my policy every night. In such a case, I don't really have a policy of eating only fruit for dessert, or if I do, I am failing to comply with it. In sum, it seems that the following argument constitutes a powerful argument against PAM:

[VIOLATION] If I had had sweet dessert on too many occasions, I would have been irrational.

[CULPRITS] If eating sweet dessert on too many occasions was irrational, then on at least some of these occasions it was irrational to eat dessert.

[GENERALITY] If eating sweet dessert is irrational on some occasion, it must be irrational on every other relevantly similar occasion.

[VIOLATION] is a straightforward consequence of the principle of instrumental reasoning and our stipulations; since we stipulated that my end is to follow the policy, violating it must be instrumentally irrational. I do not want to quibble with [GENERALITY]; some fans of extreme versions of particularism might take issue with such an assumption, but I find [GENERALITY] overwhelmingly plausible. What about [CULPRITS]? [CULPRITS] essentially claims that if a certain extended pattern of activity was irrational, there must have been momentary actions that were irrational. If we accept PAM's identification of policies and actions, [CULPRITS] would say that there could be no extended action that is irrational without some of its momentary parts being irrational; the rationality of acting in an interval supervenes on the rationality of the momentary actions in the interval. In other

words, [CULPRITS] is the denial of the possibility of what I'll call "top-down independence".⁵⁸ The rationality of a long term activity A is top-down independent if and only if there is a set of possible choices S for the agent such that (relative to A):

- (i) No particular choice in S would be irrational
- (ii) If the agent were to make all the choices in S, she would be for that very reason irrational.

If top-down independence is possible, then there can be a *pattern* of activity that is irrational relative to the policy without any of the momentary actions that constitute the pattern being irrational. Suppose, for instance, I have a policy of not drinking too much at parties. We might ask what would count as a successful execution of this policy. It seems plausible that there is no precise set of momentary choices that constitute *the exact maximum* amount of drinking allowed in a party.⁵⁹ Given the vague nature of what counts as success in the policy, there'll be clearly acceptable choices, clearly unacceptable choices, and a certain penumbral area. Because of the penumbral area, there'll be no point at which it would be right to say that if I were to drink one more drop of alcohol I will have at this point violated my policy. So in this case, at each momentary choice node, there might be no decisive reason for me not to block the policy; but if I always choose to block the policy I will fail to execute it. Having a policy with such a vague end is not essentially different from having a vague policy as defined earlier; in fact, there is no important difference between this policy and a policy not to drink at parties that allows for some unspecified exceptions.

If we were antecedently persuaded that top-down independence is impossible, and if we were convinced that a pattern of activity could be irrational only if an agent made some irrational momentary choices, then the existence of such patterns would be a powerful motivation to accept TTM Strong. For a rational agent, the successful execution of the policy is guaranteed by the existence of dispositions that pass muster in the first tier of evaluation.⁶⁰ If we have good reason to reject the possibility of instances of genuinely top-down independent policies then TTM will offer us a way to ensure that the irrationality of reconsidering or blocking the policy too much supervenes on the manifestation in particular choices of irrational dispositions of reflection and reconsideration.⁶¹ On the other hand, if there are independent reasons to accept the possibility of top-down independence, then TTM loses this possible advantage over PAM.

Let us first note that the plausibility of (SUFFICIENCY_P) on its own suggests that manifesting such habits and dispositions is only *one* way in which an agent can be rational relative to a vague policy. If the agent implements the policy in an acceptable manner, and if she does so not by chance but through her knowledge of how to realize her end, then she is complying with any plausible instrumental requirement, whether or not she always manifests commendable habits or dispositions. She is, after all, non-accidentally taking sufficient means to her end of implementing the policy.

The structural parallel between (SUFFICIENCY) and (SUFFICIENCY_P) suggests stronger grounds to accept the possibility of top-down independent policies. After all, baking a cake is also a vague end; it is unlikely that there is a precise set of “distractions” such that even one more second of listening to the news would prevent me from baking a cake. The predicate “x is a cake” is probably vague, but this is not the main reason why my end here is vague or indeterminate. Awful-tasting cakes, cakes that weren’t ready till the party was over, will not count as realizing my end of baking a cake, but there is no precise last second or minimum “tastiness” that my end specifies. [SUFFICIENCY] here seems extremely plausible; the distractions in the process of baking the cake are unlikely to be always manifestations of rational dispositions. In other words, the proposal that we could not judge an agent who successfully bakes a cake in ordinary circumstances to be rational unless we can determine first that all her distractions must have been the manifestation of the right kind of dispositions seems hard to defend unless we trivialize the notion of “manifestation of the right kind of disposition”. As long as I ended up with an acceptable cake in a non-deviant and not too lucky manner, my baking of the cake did not display any instrumental irrationality in relation to this end. Why would instrumental rationality require more from us than the non-accidental successful completion of my action of baking the cake? In fact, if [SUFFICIENCY] holds for an action of mine then I have satisfied the means-ends principle in relation to the end of my action.

But once we accept [SUFFICIENCY], we are already committed to the possibility of top-down independence; in fact, if [SUFFICIENCY] is true, top-down independence is a feature of almost any extended action or pursuit. Let us look at a case in which I bake an acceptable cake through a series of interruptions that were not all manifestations of particular rational dispositions. By [SUFFICIENCY] all my actions (while I was baking the cake) are instrumentally rational in relation to the end of baking the cake; in particular, none of the “interruptions” are deemed irrational by the theory. But now compare this case with a case where I failed to bake a cake. Let us stipulate that in such a case there are *more* interruptions, but that there are no relevant differences between the added interruptions and the original ones. Thus, by [GENERALITY], we must conclude that the new interruptions are also rational actions. But, assuming that I kept the end of baking the cake throughout this time, we must conclude that the general pattern of activity is irrational relative to the end of baking the cake. After all, I did fail to bake my cake even though the necessary means were fully under my control, and there were no unexpected circumstances; all the interruptions were due to my own agency. So if [SUFFICIENCY] is correct, we can easily build a case in which we must reject [CULPRITS]; [SUFFICIENCY] thus already implies the possibility of top-down independence.

There is one way to try to block this argument: namely, one might deny the coherence of our stipulations that there are no relevant differences among all interruptions in the case in which I failed to bake a cake. Of course, it is easy to imagine that each extra time the cat is in the same position, I am just as bored, the cake is at a relevantly similar stage, etc. But, one might argue, at least some interruptions

will be importantly different; some of them must be such that, given that so many interruptions had occurred, they would prevent the cake from being baked in an acceptable manner. In other words, there must have been some momentary action *MA* such that before I performed *MA*, it was still possible for me to bake an acceptable cake, but after I performed *MA*, it was no longer possible for me to bake an acceptable cake. In that case, *MA* is our culprit; *MA* is the irrational momentary action (or at least one of the irrational momentary actions) I performed in the case where I failed to bake a cake. However, this response fails to appreciate the vague or indeterminate nature of the end of baking a cake. There might not be any precise moment such that, for instance, looking out of the window for even one more moment will suffice to prevent me from baking the cake, even if it is true that if I look out of the window for too long I will not be able to bake the cake. And even if one thinks there must be such a moment, it is not clear that it's even in principle possible for an agent to be aware of its location in time; that is, even if one insists that there must be such a moment, it's not clear that this fact would have any significance to the assessment of the subjective rationality of the agent.

Of course, there is much more to be said about the relevance of the literature on vagueness to understanding this kind of general pattern of activity that I have been calling a "vague policy". I think that no plausible account of vagueness would block the argument above, and, in fact, I think that there are cases in which [SUFFICIENCY] implies the possibility of top-down independence that do not depend on anything like a Sorities series (in particular, cases that involve repeated gambles). But this is not an argument that I can make in a few paragraphs.⁶² Fortunately, I do not need to make this argument. Showing that vague policies are incoherent or impossible, or showing that they collapse into strict or precise policies, amounts to denying the existence of the pattern identified at the end of section three. But, as we argued above, the existence of such pattern was the best chance to establish the intentional autonomy thesis; if there were no such patterns, the arguments of the previous sections would suffice to repudiate the intentional autonomy thesis.⁶³ The cost of accepting this response is to undermine the case for TTM; in other words, TTM and PAM share the assumption that rational agents might have ends, projects, plans, etc. that have the pattern described in section three; that is, a pattern such that (i) it is unproblematic to perform actions of a certain type *at each relevant choice situation*, and yet (ii) performing actions of this type *at too many occasions* would be disastrous. But, as we argued above, this pattern provides the best case for *Intentional Autonomy*; if there were no such pattern, the arguments of the previous sections would suffice to repudiate the intentional autonomy thesis.

Moreover, once we appreciate the incompatibility of [SUFFICIENCY] and [CULPRITS], we can hijack Gauthier's argument to provide another argument for the possibility of top-down independent policies. Suppose there is a set of choice situations such that each member of the set *S* includes as a possible choice a token of an action type *T* within the range of a series of alternative policies P_0 through P_n , such that P_0 calls for never performing actions of type *T*, and P_n calls for

always performing actions of type T. Now let us say that an agent facing this choice scenario knows of no relevant or salient features⁶⁴ that distinguish any particular choice situations in S from all the other choice situations in S. Now there would be no specific habits or dispositions that could be formed that would guarantee that the agent would choose an acceptable set of choices. In fact, if there are really no salient features of the situation that are accessible to the agent, and the choice situations in S could extend indefinitely through time, the only more specific habits or dispositions that the agent could form would amount simply either to never choosing actions of type T or to always choosing them. It would be implausible to think that rationality demands in that case that one either choose P_0 or P_n , but nothing in between; that is, that rationality requires that one either be a teetotaler or a complete drunk, and that it allows nothing in moderation.

6. PAM vs TTM Weak

TTM Weak claims that conformity to reasonable dispositions is permissible but not mandatory; in other words, the two tiers deliver verdicts of “rational action” but not of “irrational action”. But here we quickly run into a dilemma. So far we said nothing about which, or even whether, certain instances of policy reconsideration are irrational according to TTM Weak. Here there are two (non-exclusive) ways in which TTM Weak could give different verdicts from PAM. On the one hand, it could be more restrictive than PAM. TTM Weak could imply that there are certain cases in which I was irrationally executing my policy P within a certain period of time, without violating the policy in this period of time. But in this case TTM Weak would still run afoul of [SUFFICIENCY_p] and our arguments against TTM Strong would still apply.

On the other hand TTM Weak could be more permissive than PAM; in this case, TTM would allow that certain cases of failing to execute a policy would count as rational that PAM would not. Perhaps reasonable habits would, given some unfortunate circumstances, lead to failure of implementing a policy. Given that so far, we only characterized PAM in terms of sufficient conditions for rational action, any further permissions that TTM Weak postulates are fully compatible with PAM. But let us look at a stronger version of PAM, a version to which we add the following principle:

[NECESSITY] In cases in which there is no risk, no relevant change of information, and in which ϕ -ing (executing a policy P) is (known to be) fully under my control, for my actions to be instrumentally rational in relation to the end of ϕ -ing (executing policy P), it is necessary that I ϕ (execute the policy) through my actions in expected ways.

[NECESSITY] is, of course, just [SUFFICIENCY] (and [SUFFICIENCY_p]) with “necessary” replacing ‘sufficient’. Thus, it does not introduce any new principles to our theory of rationality. It is rather a closure condition for the instrumental rationality of action under certainty or knowledge; it simply says that there are no

principles other than the principle of instrumental reasoning that could also make an agent instrumentally rational in relation to a policy under these conditions.

Accepting [NECESSITY] opens up the possibility that versions of TTM Weak could classify as reasonable failures some cases that PAM classifies as cases of irrationality. Is it plausible to suppose that, in cases in which agents know the possible outcomes for each occasion on which they are presented with a momentary choice, they could still reasonably fail to implement their policy adequately? Could someone rationally fail to drink moderately even if on each occasion she knew full well what the outcome of her choices would be? Perhaps we can imagine a situation like the following: Mary has reasonable habits of reconsideration; for instance, she only reconsiders her policy not to drink if there is some special circumstance in which drinking would be especially appropriate. For Mary, weddings are “relevantly special”; these are unique circumstances in which we can express our sympathetic joy by heartily toasting to brides and grooms. Unfortunately, weddings keep coming up, and Mary’s actions end up amounting to a clear failure to bring about her end. Should we conclude that Mary’s failure to keep up her policy was not a failure of rationality? If Mary is going to a wedding every day, she must realize that her dispositions of reconsideration are leading her astray; the fact that these dispositions are rational in other contexts cannot be an excuse if they are known to be inadequate in the present context. In fact, we had seen earlier that TTM, or at least TTM Strong, could be too demanding. Now we notice for similar reasons that it can be too lax; it might permit excessive reconsideration. It sounds like a joke to say in this situation that Mary was unlucky: “she has such reasonable dispositions, and she is overall such a rational agent, but, alas, there were just too many weddings”.

It might seem that we’re ignoring the obvious. If in the current context the disposition is leading Mary astray, isn’t this be enough to show that the disposition is not a rational one? Wouldn’t all rational dispositions be sensitive to the presence of too many weddings? Indeed, if we say that the correct dispositions are those which ensure that the agent complies with her policies, then TTM Weak will deliver the same verdicts as [NECESSITY]. However, [NECESSITY] allows us to get these results without positing irreducible norms governing FDIs.

One might complain that PAM’s savings are merely apparent. It cuts down on rational requirements only by inflating what goes into the traditional conception of the principle of instrumental reasoning; after all, [SUFFICIENCY_p] is not part of the traditional conception of instrumental rationality. However, strictly speaking, [SUFFICIENCY_p] is just an instance of [SUFFICIENCY] applied to policies; the more general principle is needed already to understand what counts as rationally pursuing the end of baking a cake. Results like the possibility of top-down independence might seem surprising, but they are already a consequence of [SUFFICIENCY]. The focus on mental attitudes rather than extended actions might have obscured this for us, but applying the principle of instrumental reasoning to vague or indeterminate pursuits (which are nearly all our pursuits) already commits us to top-down independence.

7. Conclusion

I started by arguing that there are no irreducible requirements concerning intention reconsideration when we look at specific FDIs. However we did find that rational agents with long-term ends should not reconsider too much, and that the requirement not to reconsider too much had a problematic structure. Given that this problematic structure was characteristic of a wide variety of general intentions, one could propose that TTM could be defended as the best theory of rationality for implementing general intentions that have this problematic structure. This proposal turned out to be equivalent to examining whether TTM was the best account of the rational pursuit of vague policies. Here I argued that PAM does a better job than TTM in determining the rational requirements generated by these policies without postulating any intention-specific requirements of nonreconsideration. I hope to have shown that PAM is, more generally, a plausible theory of the rationality of long-term policies, projects, etc. However, since PAM claims that long-term projects, policies, etc. are simply instances of extended action, PAM is best understood as an application of a very minimal theory of instrumental rationality applied to activity that extends through time to a particular case.

It has been a general assumption of theories of instrumental rationality that their principles must apply primarily to mental states, and only derivatively to actions. I think Bratman, among others, correctly noticed that principles governing the agent's belief, desires, or preferences could not exhaust the theory of instrumental rationality; he went astray, however, in thinking that the fix was to add one more mental state to the pile: future-directed intentions. A theory of instrumental rationality is a theory of rationality in action, and thus we need to turn our attention to the actions themselves. Establishing this general view is a more ambitious project; here I can claim to have established only a much more modest result. I have proposed that if we focus on how the principle of instrumental reasoning applies to extended actions with vague or indeterminate ends (and how many actions are punctate; how many of our ends are fully determinate?), we can see how policies, plans, projects, etc. are just instances of actions (at least for the purposes of the theory of instrumental rationality) and how their rational structure can be accounted for by the same meager resources; namely, the principle of instrumental reasoning.

Notes

¹ Unless otherwise indicated, I will use "action" always to mean "intentional action".

² Michael Bratman (2012, p. 79).

³ Although the principle makes mention of an interval, the interval simply guarantees the absence of other mental states that might undermine the rational relation between the mental states at these two different moments.

⁴ See McClennen (1990). For related principles, see Gauthier (1997).

⁵ For other pieces of the puzzle, see Tenenbaum & Raffman (2012), Tenenbaum (2010), Tenenbaum (2014), and Tenenbaum (2015).

⁶ I look at other norms in Tenenbaum (2014).

⁷ This assumption is not strictly necessary for my argument; all we need to assume is that the intention is based on reasons. But it'll make it for a simpler presentation to keep the stronger assumption in place.

⁸ I am largely in agreement with Luca Ferrero's view that what he calls "the division of deliberative labour" does not threaten the autonomy of my future self (See Ferrero (2010)). However, as will be clear later, I do not think that the authority of my past intentions should be understood in terms of exclusionary or protected reasons. For a more detailed criticism of Ferrero's understanding of the authority of my past intentions, see Nefsky & Tenenbaum (n.d.).

⁹ This example is adapted, of course, from Parfit (1984).

¹⁰ I formulate the case in terms of preference shift but it's easy to reformulate in terms of a shift in evaluative judgment.

¹¹ Kavka (1983).

¹² Holton (2004, 2009).

¹³ More specifically, in order for it to apply to cases like (A), the choice in (A) must be optimal given the preferences of my earlier and later self. We'd need to modify the case slightly by, for instance, assuming that I'd think it is worth selling my company earlier and make less money if I know that I would not donate half my money to charity if I were to become a millionaire.

¹⁴ Gauthier (1997), McClennen (1990). In fairness to Gauthier and McClennen, their concerns for optimality are not necessarily best understood as an "extension". of the consideration in [SETTLE] and [LONG].

¹⁵ Bratman (2007).

¹⁶ Gauthier gives the example of a boy who finds girls "yucky", forms the intention now not to date them later as he suspects that he'll not have the same attitudes in the future. It seems absurd to think that this earlier intention should require anything from his later self. See Gauthier (1997).

¹⁷ George Ainslie has famously noted the ubiquity of hyperbolic discounting in human agency in Ainslie (2001). Holton relies on these and related phenomena in his arguments in favour of rational resolution. See Holton (2004, 2009).

¹⁸ Bratman (1999, p. 82).

¹⁹ In fact, the reasons in question would be the same if instead of forming the intention to go to Paris I simply won these things in a lottery, or at least if I had inadvertently bought tickets and made the reservations.

²⁰ There are complications that can be introduced here. For instance, someone might have poor habits of nonreconsideration, but this would not necessarily mean that she is never rational when she reconsiders, or fails to reconsider, an intention. Some of these complications will be important later when I distinguish various versions of the two tier model, but for now this simplified formulation is good enough for our purposes.

²¹ Means-ends coherence principles are possibly as important. But, it is not clear that a principle of means-end coherence for FDI's in particular is indispensable. Once we apply principles of coherence governing beliefs to beliefs that are formed in forming FDI's (such as the belief that I will ϕ that typically results from forming an intention to ϕ), and transmission principles governing normative reasons, we might capture all that a principle of means-end coherence would capture when applied to FDI's. However, I cannot go into this issue in detail here (but see Tenenbaum (2014)). For reasons that will be clear below, I do not think we could dispense with a principle of instrumental reasoning that applies to *ends that are currently being pursued*.

²² Bratman has lately defended the view that diachronic requirements on intention can be grounded on considerations of self-governance (see, for instance, Bratman (2009, 2012)). I can't discuss these topics in detail here; however, I hope it will be reasonably clear how the arguments in the paper would extend to the considerations of self-governance. For more detailed discussion of Bratman's arguments in these papers, see Nefsky & Tenenbaum (n.d.), Paul (2014).

²³ One might complain that one of the things we want to know is whether we are now facing exigent circumstances or not, and it would be disastrous to deliberate if this were the case. But on any view there must be salient features of exigent circumstances (otherwise it could not trigger a habit of non-reconsideration in exactly these circumstances), and it is enough to postulate that one has a reason (or is under a requirement) not to reconsider under such circumstances. These reasons or requirements will "kick in" whether or not one has formed an intention.

²⁴ For a similar point, see Parfit (2011).

²⁵ Of course, there might be requirements not to revise your justified doxastic states without any change in evidence (at least under certain conditions), but a requirement *not to reconsider any belief* without a change in evidence would be too strong (if, for instance, I think that my current belief that *p* might not be justified in light of the evidence I have, it would be, *ceteris paribus*, permissible to reconsider it even if I have no new evidence).

²⁶ And if one's conceptions of desire, hope, trust, etc. allow for reconsidering one's desire, hope, trust, etc., then similar norms will apply to reconsideration of such attitudes.

²⁷ Bratman claims that intentions resist reconsideration in Bratman (1999). I discuss below various normative principles related to this idea. Bratman also defends a wide scope requirement to persist in one's intention through time (under certain conditions) in Bratman (2012).

²⁸ John Broome has recently defended some intention- and belief- persistence requirements (Broome, 2013). I think they are equally problematic, but I cannot discuss them here. See Tenenbaum (2014) for discussion of these requirements.

²⁹ Including, of course, a second order end to realize larger ends (the end of having a career, to engage in meaningful activities), etc.

³⁰ We discuss the rational structure of such cases also in Tenenbaum & Raffman (2012).

³¹ For a more detailed discussion of how to understand cases in which various looser policies are not feasible for the agent, see Tenenbaum (2010). One might think that given that there is no principled way to determine when we should make exceptions to the policy or even precisely how many exceptions to the policy are permissible, the only possible rational policy is a strict resolution. I discuss this type of argument in section five.

³² For a sustained defence of such intention-specific requirements, see Holton (2009). I am raising some doubts about the conclusions of this book, but I can't do justice to Holton's detailed argument there. For more detailed criticism of Holton's views on these issues, see Paul (2011).

³³ This is a bit simplified. There are other things I might be doing while swimming (I might be checking around for sharks, fixing my hair, etc.). It would be more precise to say that whenever I am acting I am doing something that is a means to the action of swimming. Moreover I am ignoring complexities introduced by the fact that the constitutive or instrumental means can themselves be gappy. So we don't want my writing the book to be gapless because I am always writing a chapter of my book while I am writing my book (given that "writing a chapter" is itself a gappy action).

³⁴ See Falvey (2000).

³⁵ For simplicity, I will talk only about "policies", but what I say about policies should generalize to long-term projects, plans, etc.

³⁶ I will be talking about momentary actions, but there is some skepticism about this notion. See Thompson, *Life and Action*. For my purposes, the notion of "momentary action" can be understood as relative to a larger action; in this context, a momentary action relative to a larger action *A* is an action that takes place during an interval that is a proper part of the interval in which *A* takes place.

³⁷ Vendler (1957). See also Mourelatos (1978). Although I am using the same labels as Vendler and Mourelatos, my usage is slightly different from theirs.

³⁸ See, for instance, Thompson (2008), and Steward (2012). These philosophers often claim that focusing on the progressive aspect of action-ascriptions, or on actions' nature as processes, has important consequences for the metaphysics of action.

³⁹ See, for instance, Moss (2015).

⁴⁰ "X-ing" here could refer here to a state (such as "believing" or "deciding") or to a "doing".

⁴¹ Typically, *t* is later.

⁴² I will argue below that this assumption does not hold for the rationality of actions. For all I say in this paper, it might be true that time-slice views and diachronic views are exhaustive in epistemology.

⁴³ Anscombe (2000, p. 55). The original example is not an illustration of instrumental irrationality but of contradictory expressions of intentions by two different persons.

⁴⁴ I am ignoring further constraints that my end might have; it might be important for me to have the cake baked by a certain time, etc. "In expected ways" is supposed to rule out deviant causal chains.

⁴⁵ Recall that we are examining only rational requirements governing (non)reconsideration.

⁴⁶ Henceforth, I'll just use "dispositions".

⁴⁷ Holton (1999, p. 248) implicitly makes a similar distinction and endorses the stronger view.

⁴⁸ There are many other possible “degrees of strength”, but this complication is irrelevant for our purposes.

⁴⁹ More precisely, we should distinguish between loose and strict, and among the loose policies, vague and non-vague. A more general notion of vague policies (or long-term ends) would allow, not only the conditions under which a policy is blocked, but also the conditions under which it is successfully executed, to be vague.

⁵⁰ “Everybody Knows” © Universal Music Publishing Group, Sony/ATV Music Publishing LLC.

⁵¹ Of course, they can also be violated. The distinction between “blocking” and “rejecting” a policy was introduced by Bratman (1999).

⁵² “Loose-and-precise” might be thought to mean the same as a strict policy. But they are not quite the same. In a strict policy, every instance of conformity is sufficiently explained by the policy. If I didn’t accept a bribe today, my strict policy never to accept bribes fully explains my not accepting today’s bribe. But an exception cannot be explained this way. My drinking beer on Friday is not sufficiently explained by my policy of drinking no more than two pints; I also need a reason to drink. However, the difference is not going to be very important for our purposes. It’s also worth mentioning that the coherence of loose but precise policies is dubious. If I add a drop to my two pints of beer, have I violated my policy? I’ll ignore these complications.

⁵³ This is another way to put what I take to be correct in Bratman’s claim that “intentions resist reconsideration”. Of course, if I am right, the language of “resisting” is highly misleading; reconsideration is just a form of interfering with a state that would otherwise persist through time.

⁵⁴ I consider the other possible divergences in the next section.

⁵⁵ At least as long as the pattern was brought about by my actions in non-deviant manner.

⁵⁶ The relevant habits and dispositions will probably be more general than this, but it simplifies matters to consider a rather specific habit.

⁵⁷ Gauthier (1997, p. 21).

⁵⁸ I define this notion in Tenenbaum (2010). See Tenenbaum & Raffman (2012) for further discussion of this notion.

⁵⁹ Not to mention that there’ll be no precise criteria for when such a policy can be blocked without being violated, no precise criteria for what counts as a party, etc.

⁶⁰ Assuming no unforeseen circumstances or bad luck.

⁶¹ Or on particular choices that fail to manifest rational dispositions.

⁶² We argue for something like this general conclusion in Tenenbaum & Raffman (2012).

⁶³ At least in the limited version we’re considering here in which the intention specific requirements govern reconsideration.

⁶⁴ Or even reliably expects that there will be no such features.

References

- Ainslie, G. (2001). *Breakdown of the Will*. Cambridge: Cambridge University Press.
- Anscombe, G. E. M. (2000). *Intention* (2nd ed.). Cambridge, MA: Harvard University Press.
- Bratman, M. (1999). *Intention, Plans, and Practical Reason*. New York: Cambridge University Press.
- Bratman, M. (2007). Temptation Revisited. In his *Faces of Intention* (pp. 257–282). Oxford: Oxford University Press.
- 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.
- Broome, J. (2013). *Rationality Through Reasoning*. Oxford: Wiley-Blackwell.
- Falvey, K. (2000). Knowledge in Intention. *Philosophical Studies*, 99(1), 21–44.
- Ferrero, L. (2010). Decisions, Diachronic Autonomy, and the Division of Deliberative Labor. *Philosophers’ Imprint*, 10(2), 1–23.
- Gauthier, D. (1997). Resolute Choice and Rational Deliberation: A Critique and a Defense. *Noûs*, 31(1), 1–25.
- Holton, R. (1999). Intention and Weakness of Will. *The Journal of Philosophy*, 96(5), 241–262.
- Holton, R. (2004). Rational Resolve. *The Philosophical Review*, 113(4), 507–535.
- Holton, R. (2009). *Willing, Wanting, Waiting*. Oxford: Clarendon Press.



- Kavka, G. S. (1983). The Toxin Puzzle. *Analysis*, 43(1), 33–36.
- McClellenn, E. (1990). *Rationality and Dynamic Choice: Foundational Explorations*. New York: Cambridge University Press.
- Moss, S. (2015). Time-Slice Epistemology and Action Under Indeterminacy. *Oxford Studies in Epistemology*, 5, 172–194.
- Mourelatos, A. P. D. (1978). Events, Processes, and States. *Linguistics and Philosophy*, 2(3), 415–434.
- Nefsky, J., & Tenenbaum, S. (n.d.). Extended Agency and the “Problem” of Diachronic Autonomy.
- Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- Parfit, D. (2011). *On What Matters*. Oxford: Oxford University Press.
- Paul, S. (2011). Review of “Willing, Wanting, and Waiting” by Richard Holton. *Mind*, 120, 889–892.
- Paul, S. (2014). Diachronic Incontinence is a Problem in Moral Philosophy. *Inquiry*, 57, 337–355.
- Steward, H. (2012). Action as Processes. *Philosophical Perspectives*, 26, 373–388.
- Tenenbaum, S. (2010). The Vice of Procrastination. In C. Andreou & M. D. White (Eds.), *The Thief of Time: Philosophical Essays on Procrastination* (pp. 130–150). New York: Oxford University Press.
- Tenenbaum, S. (2014). Minimalism about Intention: A Modest Defense. *Inquiry*, 1–28.
- Tenenbaum, S. (2015). Acting and Satisficing. In G. Pavlakos & V. Rodriguez-Blanco (Eds.), *Reasons and Intentions in Law and Practical Agency* (pp. 31–51). Cambridge: Cambridge University Press.
- Tenenbaum, S., & 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*. Cambridge, MA: Harvard University Press.
- Vendler, Z. (1957). Verbs and Times. *The Philosophical Review*, 66(2), 143–160.