Plan Rationality[1]

Michael E. Bratman   01-25-19

Forthcoming (with possible small changes) in *The Routledge Handbook of Practical Reason,* edited by Ruth Chang and Kurt Sylvan. Please quote from published version.

Plan rationality?

An important idea in practical philosophy is that there are norms of rationality that apply to our intentions. (Harman 1976; Bratman 1987) One example is a consistency norm according to which both intending A while believing not-A, and intending A and intending B while believing that A and B are not co-possible, are rational breakdowns.[2] A second example is a norm that enjoins coherence between intended ends and intended means: there seems to be a rational breakdown in intending E, believing that a necessary means to E is M and that one will M only if one now intends M, and yet still failing to intend M.

There are debates about how exactly to understand such norms. Perhaps they articulate local or pro tanto, rather than global or on-balance rational pressures. (Smith 2004) Perhaps one might, for example, achieve local, but not global, rationality in intending known necessary means to an intended end even though the intention in favor of the end is not sufficiently supported by what one sees as one's relevant reasons. Further, perhaps there are special circumstances in which these norms do not apply in the standard way. An example might be a planning analogue of the paradox of the preface in which the consistency norm does not apply in the standard way. (Bratman 2009a: 413; Shpall 2016; Goldstein 2016) Finally, there are issues about scope: does the means-end coherence norm have "narrow" scope and say, roughly, that if you intend E and believe M is a necessary means to E then there is rational pressure to intend M? (Kolodny 2005) Or does it rather have "wide" scope and say that there is rational pressure, roughly, to be such that if you intend E and believe M is a necessary means to E then you intend M? (Broome 2013)

Here I will suppose that these are norms of local or pro tanto rationality.[3] Further, I will suppose that these are wide scope norms that enjoin a certain organization of one's mind. One can, for example, avoid the threat of means-end incoherence by intending the cited means; but one can also avoid it by giving up the intended end. So, there is here a normative symmetry, though it remains possible that other rational pressures (e.g., pressures for intention stability) asymmetrically favor one of these ways of avoiding incoherence.

These norms of consistency and means-end coherence of intentions directly concern the organization of one's intentions at a given time (or during a small temporal interval), though one may well adjust one's intentions over time (e.g., by adding an intention in favor of necessary means) in order to avoid violating these norms. There is also room for norms that directly concern the development over time of one's system of intentions. An example might be a norm that favors a default persistence of prior intention. I will turn to this idea below. For now let's focus on the cited synchronic norms.

Consider two reasons why such norms can seem puzzling. First, there are other practical attitudes with respect to which it is less plausible that analogous norms apply. Consider ordinary desires. It is a common feature of our lives that we desire things we know not to be co-possible. And it is unclear why there is any rational breakdown in desiring, for example, to be very wealthy while finding unattractive, and so not desiring, what one thinks are necessary means. If our thoughts about synchronic rationality of intention are on the right track there must be something about intention that distinguishes it from ordinary desires and helps explain these differences in relevant rationality norms.

Second, when we reflect on these wide-scope norms they can seem simply to be recommending a kind of mental tidiness. But why should such tidiness be a big deal? (Raz 2005; Kolodny 2008; McCann 1991)

To respond to these challenges we need a systematic account of intention itself: there is here a tight connection between questions in normative philosophy about practical rationality and questions in the philosophy of action about the role of intention in our agency.

My proposal highlights an apparently basic feature of our human agency: we are planning agents. Given needs for coordination over time and with each other, and given our resource limits (Simon 1983), we normally settle in advance on prior partial plans and then fill them in as need be and as time goes by. Not all purposive agents are planning agents: there can be purposive agents – sheep, perhaps -- who act in order to pursue their goals, and in light of their cognitive grasp of the world, but who are not planning agents.[4] In contrast, we are planning agents, and this is a basic element in our capacities for intra-personal and inter-personal organization of our agency. (Bratman 1987, 2010, 2014) And my proposal is that we understand the intentions that are the target of the cited rationality norms as states in this planning system: intentions are plan states.

To spell this out we need a model of the normal functioning of such a planning system. Here a basic idea is that this functioning involves guidance by the cited norms. (Gibbard 1990; Hart 2012) Given our resource limits, our prior plans will exhibit a characteristic partiality. Given this partiality, it is because of pressures for means-end coherence that one normally is led to questions of means and preliminary steps. And in solving these problems of means-end coherence one is constrained by pressures of consistency of intention. In this normal functioning of the planning system one's practical thinking is guided by one's at least implicit acceptance of norms of plan coherence and consistency.

Within this model these norms apply to intentions in a way in which they do not apply to ordinary desires. But how we can justify this distinctive normative guidance?

An initial proposal is that guidance by pressures for plan consistency and coherence supports important forms of intra-personal and inter-personal coordination and organization. Intending what one takes to be incompatible options would normally lead one to trip over oneself, and intending ends without intending believed necessary means would normally block the efficacy of one's planning system. At least, this is true if one's relevant beliefs are reliable. And in each case we should expect analogous impacts on relevant inter-personal coordination. It is in part because the planning system is guided by these norms that it tends to be successful in these coordination roles. (Bratman 1987; Morton 2017)

However, while this is true and important, it does not suffice. Perhaps in a special case one might better achieve one's ends by way of intentions and plans that are, given one's beliefs, inconsistent.[5] Or perhaps in another special case one would be rewarded for means-end incoherent plans. Inconsistency or incoherence of intention in such cases still seems a kind of rational breakdown. But why? Appeal to the coordination benefits of the general system of plan-infused practical thinking provides a defense for this *general* mode of thinking; but it is not clear how to get from a defense of this general mode of thinking to a defense of its application in each *particular* case. (Smart 1956)

One response would turn to theoretical rationality. If intentions involve associated beliefs then perhaps we can see norms of plan consistency and coherence as riding piggy-back on theoretical norms of consistency and coherence of belief. (Harman 1976; Setiya 2007; Wallace 2001; Velleman 2014) This is *cognitivism* about these norms of plan rationality. (Bratman 1999a) Such cogntivism does not yet justify the cited theoretical norms. But it is sometimes progress to reduce two problems to one.

How might this work? Suppose – for the sake of argument -- that intending that p necessarily involves believing that p.[6] We could then argue that inconsistent intentions

4

necessarily involve inconsistent beliefs. And it might seem that we could also argue that means-end incoherence of intention necessarily involves incoherence of belief. After all, if you intend E, and so (on present assumptions) believe E, and if you believe that E only if M and that M only if you intend M, and yet you do not intend M, then it may seem that you will believe that E without believing that a necessary condition for E will obtain. And that would be a kind of theoretical incoherence of belief.

But I do not think that this works. (Bratman 2009b) A basic problem is that one might believe that one intends M and yet not in fact intend M. In such a case one's beliefs might be coherent even though one's plans remain means-end incoherent. Granted, in such a case one has a false belief about oneself; but false belief need not be irrationality or incoherence. So there can be means-end incoherence in one's plans without incoherence in one's beliefs.[7] So this cognitivist strategy does not get to the bottom of why means-end coherence of plans matters.

Another strategy might be to highlight a kind of inescapability.[8] After all, we have supposed that guidance by these norms is built into our planning agency. But the problem here is that, as noted, there are multiple forms of agency; in particular, one can be an end-directed, purposive but non-planning agent. So, guidance by these norms is not inescapable for agency. So, it is not clear how the idea that guidance by these norms is partly constitutive of, in particular, planning agency can explain why these are norms of practical rationality.[9]

Donald Davidson highlights a different kind of inescapability. Treating norms of rationality as a single over-all package, Davidson writes that "it is only by interpreting a creature as largely in accord with these principles that we can intelligibly attribute propositional attitudes to it …An agent cannot fail to comport most of the time with the basic norms of rationality." (Davidson 2004) This is a plausible idea about interpretation; but it does not solve our problem about plan rationality. What is claimed to be inescapable for a person with a mind is only failing

to "comport most of the time" with relevant norms. (Kolodny 2008) So we do not have an explanation of why a violation of these norms is, in each particular case, a rational breakdown. Further, Davidson is here treating norms of rationality as a single package, one involved quite generally in interpreting minds. But given the multiplicity of agency there can be minded agents who are not planning agents and so are not subject to norms of, in particular, plan rationality.

What do we need to show to defend the claim that these are norms of practical rationality for a planning agent? Suppose you are – as there is pragmatic reason to be -- a planning agent. Your practical thinking involves guidance by the cited norms. You can nevertheless step back and ask whether this structure of thinking makes sense in a way that supports its stability under your reflection. And your answer will have implications for our descriptive and explanatory model of human agency. After all, our confidence in the basic descriptive and explanatory significance of planning structures in human agency would be to some extent challenged if these structures would not themselves be stable under a planning agent's reflection.

So, let's focus on you – a planning agent reflecting on her characteristic forms of practical thinking and asking whether they make sense. Part of your answer will return you to the pragmatic benefits – especially given your resource limits and needs for coordination, both cross-temporal and social – of planning agency. But you will also need to supplement these observations in order fully to defend the application of these norms to each particular case. How?

The strategy of self-governance

Here I think we make progress by understanding the ways in which these norms track central conditions of a planning agent's self-governance. Call this the *strategy of self-governance*. (Bratman 2009a, 2010, 2017, 2018a) Successfully to pursue this strategy you need

to do two things: (1) Articulate ways in which these norms track conditions of a planning agent's self-governance. (2) Explain how this provides further support for these norms.

Beginning with (1), we can draw on the Frankfurt-inspired idea that an adequate model of self-governance at a time (or during a small temporal interval) involves the idea of "where (if anywhere) the person himself stands." (Frankfurt 1988: 166). In self-governance one's thought and action are guided by where one relevantly stands; and one's standpoint is constituted by a web of relevant attitudes, where this web (as we can say) speaks for the agent. To play this role this attitudinal web will need to be sufficiently coherent (Bratman 2009a) – though such coherence does not require that the agent have, in particular, the end of synchronic self-governance. (Bratman 2018b: 243) And the idea is that guidance by such a coherent web of *attitudes* can constitute the *agent's* self-governance.

The next idea is that, given the basic roles plan states play in your practical thought and action, your relevant attitudinal webs will be plan-infused. In planning one is settled on certain courses of action and/or the relevance of certain considerations to one's ongoing practical thinking.[10] One's plan states will normally cross-refer: one's plans for the future will at least potentially refer to one's future intentions which will normally refer back to one's earlier plans. Further, there will normally be a kind of interdependence between, on the one hand, intentions at a given time and, on the other hand, past intentions and expectations of future intentions. These cross-referring, issue-settling, interdependent plans will frame one's on-going thought and action in ways that tend to track and to support mesh between sub-plans at different times. In these multiple ways these plan states will normally induce forms of psychological continuity and referential connectedness familiar from Lockean models of personal identity. (Parfit 1984; Yaffe 2000; Bratman 2000)

This supports the conclusion that a planning agent's self-governance at a time will involve relevant consistency and means-end coherence of plan. After all, if you intend A and intend B, but believe that A and B are not co-possible, there will be no clear answer to the question of where you stand with respect to these options. And if you intend E but nevertheless fail to intend known necessary means, even though you know that such an intention is now needed if you are to pursue those necessary means, then there will be no clear answer to the question of where you stand with respect to E. So we can conclude that the cited norms of plan rationality do indeed track basic conditions of a planning agent's synchronic self-governance.

But how does this connection to self-governance provide the support for these synchronic planning norms that our reflective planning agent is seeking, support that supplements the pragmatic reasons that favor relevant general forms of thinking?

An initial answer has two prongs. First, this connection between synchronic planning norms and synchronic self-governance reveals an overarching order within, and commonality across, those planning norms. These norms do not just track disparate forms of mental tidiness: they track conditions of self-governance. And this overarching commonality helps make sense of these norms.

The second prong involves the idea of a normative practical reason. Suppose one has a normative reason in favor of one's self-governance. And suppose one has the capacity for relevant self-governance. Given the way in which the cited norms track necessary conditions of a planning agent's self-governance, we can plausibly infer that one will have a normative reason of self-governance to conform to these norms.[11] And the idea is that appeal to this reason is part of a two-pronged argument, available to a reflective planning agent, that in tandem with pragmatic support for general structures of plan-infused thinking supports the conclusion that these are indeed norms of practical rationality for her.[12]

To fill this in we will need to say more about the very idea of a normative practical reason. But first let's ask how this approach to synchronic plan rationality might bear on the question of whether there is, as well, a norm of diachronic plan rationality. Can we extend the strategy of self-governance in the direction of a norm concerning the stability of intention?

Extending the strategy of self-governance: 'acting together' with oneself over time

In the background are considerations concerning intention stability other than ones that appeal to self-governance. There may well be pragmatic grounds for a general tendency toward sticking with one's prior intentions. After all, this tendency may make one both more likely to resist sudden and temporary shifts in desire and a more reliable partner in joint activities. But this is at most an argument for a general disposition of thought and action; and there will in any case be a presumption against stability of prior intention in the face of new information that the agent sees as undermining the case for the intended action.

To this we can add two ideas. (Bratman 1987) First, a prior intention at t1 to A at t2 may well lead to action that changes the downstream circumstances in ways that help support A-ing at t2. This is the snowball effect. Second – and especially for resource-limited agents like us -- reconsideration of a previously formed intention involves characteristic costs and risks to previously forged coordination. So, in many cases it will be sensible not to reconsider. And such sensible non-reconsideration will support the stability of one's prior intention.

Against this background, the extended self-governance strategy helps us articulate a further consideration in favor of stability of intention. It does this by way of appealing not only to self-governance *at* a time but also to self-governance *over* time.

What is a planning agent's self-governance *over* time? My proposal is that a planning agent's self-governance over time involves her self-governance at times along the way together with relevant cross-temporal inter-connections between these forms of synchronic self-

governance. (Though see (Nefsky and Tenenbaum forthcoming).) What inter-connections? An initial appeal is to the cross-temporal inter-relations between planning attitudes, inter-relations that are characteristic of planned temporally extended activity. As noted, these will include continuity of intentions over time, cross-references between intentions at different times, intended mesh between sub-plans at different times, and characteristic forms of interdependence. Further, these cross-temporal *intra*-personal continuities, interconnections, inter-dependencies, and intended forms of mesh between plan states over time will be analogous to the *inter*-personal inter-connections of plan states that are characteristic of inter-personal shared intentional activity, as understood in (Bratman 2014). (Bratman 2018b) These observations support the metaphor that in diachronic self-governance a planning agent is 'acting together' with herself over time. On this approach, then, a planning agent's self-governance over time involves her self-governance at times along the way together with forms of intra-personal diachronic intention continuity, inter-connectedness, interdependence, and intended mesh that are analogous to inter-personal inter-relations characteristic of shared agency.[13] [14]

If that is how we understand a planning agent's self-governance over time, might we see diachronic plan rationality as, in part, tracking such diachronic self-governance? Well, if it did then we would have the basis for a (perhaps minimal) presumption in favor of continuity of intention over time, so long as that continuity coheres with self-governance at times along the way. This would be a kind of conservatism, one that gives a prior intention a (perhaps minimal) default status.[15] But it would be only a modest conservatism since this default status would be present only so long as following through with one's prior intention itself cohered with synchronic self-governance at the time of follow through.

Shuffling

Consider now a case in which one decides on a temporally extended option in the face of what one sees as non-comparable considerations in favor of conflicting options. (Broome 2001; Bratman 2012) In a version of Sartre's example, one decides in favor of staying with one's mother rather than fighting with the Free French. (Sartre 1975) Given that the relevant activity is extended over time, the non-comparability will normally remain as time goes by. Is there nevertheless rational pressure to stick with one's earlier decision rather than to shuffle to a decision in favor of the alternative?

The proposed connection between diachronic plan rationality and diachronic self-governance points to an affirmative answer. Given that the non-comparability continues to be recognized by you, sticking with your prior intention to stay with mother, and switching to an intention in favor of the Free French, would each, taken separately, cohere with your then present practical standpoint. But if you stick with your prior intention your relevant intentions over time will exhibit a constancy characteristic of a planning agent's self-governance over time; if, in contrast, you change your mind in favor of the Free French there will be a kind of intention discontinuity that is in tension with self-governance over time. So, conditions of diachronic self-governance favor the stability of your prior intention. So, if diachronic plan rationality tracks conditions of diachronic self-governance it will in this respect favor this stability.

Temptation

Consider now a puzzle about rational willpower. (Holton 2009; Bratman 1999b) Suppose you know you will be tempted to drink a lot at tonight's party. Since you now think that would be a mistake, you now form the intention to stick with one drink at the party. You know, however, that at the party your judgment will temporarily shift and you will at least initially judge that it would be better to have many drinks – though you also know that if you did give into this temptation you would later regret that. If at the party you were to follow through with

your prior intention you would be acting contrary to your then-present evaluation. So how could it be rational for you to follow through on your prior intention?

It might seem that our comments about a planning agent's diachronic self-governance can help. If you stick with your prior intention there will be an important continuity in intention; in contrast, if you give into temptation there will be a break in this continuity. So, if diachronic plan rationality tracks central conditions of a planning agent's diachronic self-governance it may seem we have a way of explaining how such willpower can sometimes be rational.

But there is a problem. Continuity of prior intention helps constitute diachronic self-governance only given synchronic self-governance at times along the way. But if your judgment shifts at the time of the party it seems that sticking with your prior intention would not be a case of synchronic self-governance: after all, given the judgment shift, where you stand at the time of the party seems to favor drinking more.

What we need is an end of the agent's that in some way favors relevant intention continuity and thereby can potentially help re-shift her standpoint at the time of the party back in favor of stopping with a single drink, as she had intended.[16] Such an end would open up the possibility that sticking with one's prior intention in a temptation case is indeed a case of synchronic self-governance and so, since it involves relevant intention continuity, a case of diachronic self-governance.

What end? A simple appeal to an end of intention continuity would face the worry that we are just appealing to a kind of diachronic mental tidiness.[17] A different approach would appeal to an intellectual end of self-understanding. (Velleman 2006: 272) But given our search for commonality in our understanding of these norms, this would lead to the cognitivism against which I have argued. So, I propose, instead, that we appeal to the end of one's diachronic self-governance. This end would potentially help support willpower in the face of temptation, since

such willpower would involve a continuity of intention that is an element in one's diachronic self-governance.[18] And given the support from this end, the agent's standpoint at the time of the party might well shift back into favoring willpower in a way needed for synchronic self-governance (and so for diachronic self-governance).

The next point is that problems of temptation, and related problems about procrastination, pervade our lives. (Paul 2014; Tenenbaum and Raffman 2012; Andreou 2012; Morton and Paul 2019) So this end of one's diachronic self-governance will be an element in central cases of the exercise of a planning agent's capacity for diachronic self-governance. Whereas synchronic self-governance does not require the end of synchronic self-governance, there is this complex pressure on diachronic self-governance to involve the end of diachronic self-governance.[19] This supports the idea that if diachronic plan rationality were to track conditions of a planning agent's diachronic self-governance it would not only favor intention continuity, in a context of synchronic self-governance at times along the way; it would also favor the presence of the end of one's diachronic self-governance. And this end, if present, can sometimes shift what is supported by the agent's standpoint at the time of plan follow-through.

So, this potential extension of the strategy of self-governance to diachronic plan rationality would support both a modest norm of default stability of prior intention and the presence of the end of one's diachronic self-governance. The idea is not that this end of diachronic self-governance is essential to agency quite generally. There are, as we have noted, multiple kinds of agents; there are particular cases of our all-too-human agency in which this end of diachronic self-governance seems not to be engaged; and even synchronically self-governed agency, taken on its own, does not require this end. The purported connection to this end of diachronic self-governance goes instead by way of a strong form of temporally extended agency, namely: diachronically self-governed agency. And this end of diachronic self-governance, if

present, would then induce an end in favor of synchronic self-governance insofar as synchronic self-governance is an element of diachronic self-governance.

A unified account

Thinking about diachronic plan rationality along these self-governance-based lines sets the stage for a uniform self-governance-based account of both the synchronic and diachronic aspects of plan rationality, one that appropriately supplements the pragmatic support for general forms of plan-infused practical thinking. Norms of synchronic and diachronic plan rationality are tied together, and thereby made more intelligible, by the ways in which they track conditions of a planning agent's self-governance.

This supports an extension of our earlier observation about the significance of a normative reason for self-governance. Suppose that one has a normative reason in favor of one's self-governance, both synchronic and diachronic. And suppose that one has the capacity for relevant self-governance. Given the way in which these norms track necessary conditions of a planning agent's self-governance, both synchronic and diachronic, we can conclude that one will have a normative reason of self-governance to conform to these norms.[20]

These norms of plan rationality have a kind of stringency. (Bratman 1987: 24) We can interpret this by appeal to the contrast between a pro tanto and a merely prima facie norm. The norms we have been defending do not just cite prima facie evidence in favor of a conclusion about on-balance irrationality, evidence that is potentially misleading. Instead, these norms each articulate a kind of pro tanto rational breakdown that would be constituted by their violation, a pro tanto breakdown that would remain even if, in special circumstances, it would make most sense, on balance, to violate the norm. And if there is indeed a reason of self-governance to conform to these norms in the particular case, this reason helps support this pro tanto demand.

But what is this idea of a normative practical reason, and will there indeed be such a reason in favor of a planning agent's self-governance?

A reason for, and end of, self-governance?

For our purposes of articulating the relevant reflections of a planning agent, it will be plausible to understand such normative reasons by appeal both to the agent's ends and to the desirability of what those ends favor. A planning agent who reflects on her own practical thinking will be interested in both what is needed to realize her ends and whether what those ends favor is desirable. This supports the idea that a consideration is, in the sense relevant to our investigation here, a normative reason for S to A only if it helps explain why S's A-ing is needed to realize relevant ends of S,[21] and only if what those ends favor is desirable. On the plausible assumption that it is desirable to govern one's own life,[22] a central question about a purported reason for self-governance will then concern the status of the end of one's self-governance.

We could side-step this question if the desirability of self-governance directly insured a normative reason for self-governance. This is, however, blocked by our understanding of the relation between normative reasons and the agent's ends. Nor can we infer from simply the desirability of self-governance that a rational agent will have the end of her self-governance. There are many good things and not enough time.

What then to say? One thought might be that this end is simply a rationally optional, contingent end. If present, it can help support the reason for self-governance for which we are looking; but there is no guarantee that it will be present. At the other extreme is the thought that this end is essential to agency.[23] The former view seems too weak for our purposes of grounding norms of plan rationality. But, for reasons noted earlier, the latter view seems too strong. Indeed, there even remains the possibility of a planning agent without this end: think of young human children. (Bratman 2018a: 16)

Is there a view in the middle? A planning agent who reflects on the basic structures of her practical thinking – and so on the cited planning norms – would recognize both that there is pragmatic support for these general, plan-infused modes of thinking and that these norms have in common the property of tracking conditions of her self-governance, both synchronic and diachronic. She would go on to see that an end that is an element in the normal exercise of her capacity for diachronic self-governance – the end of her diachronic self-governance – would, if present, also induce an end of synchronic self-governance. If present, this combined end of self-governance at a time and over time would ground a normative reason for both diachronic and synchronic self-governance; and this reason would support the application of her norms of plan rationality to the particular case. Since the norm of diachronic plan rationality that would thereby be supported by this end of self-governance itself supports the presence of this end, this end of self-governance would be, if present within this planning framework, *rationally self-supporting*. Given this rationally self-supporting end – an end that favors a central, organizing commonality across her planning norms -- her package of pragmatically supported plan structures and ends would be reflectively stable: this would be a *rationally stable reflective equilibrium*. So, given her end of self-governance, as well as basic pragmatic pressures, it would make sense for her to retain her plan-infused practical thinking and its associated norms, norms that support that end of self-governance. These norms would thereby have for her a thoughtful and rational stability, one that depended *inter alia* on the rationally self-supporting presence of the end of her self-governance.

The significance of a stable, rational equilibrium?

This leads to a final question: Does it suffice for our supplement to a pragmatic defense of these planning norms to show that their acceptance is an element in a stable rational equilibrium that would be characteristic of a diachronically self-governing planning agent? Is

this a philosophically adequate path between the Scylla of a merely contingent, rationally optional end of one's self-governance and the Charybdis of an insistence that this end is essential to agency? Our investigation into norms of plan rationality has led us to this general question about the philosophical significance of fecund structures of practical thinking that are in a rationally stable reflective equilibrium though they are not strictly necessary for agency, or even planning agency, *per se*.[24]

Michael E. Bratman is U. G. and Abbie Birch Durfee Professor in the School of Humanities and Sciences, and Professor of Philosophy at Stanford University, U.S.A. His most recent book is *Planning, Time, and Self-Governance: Essays in Practical Rationality* (2018).

References

Andreou, C. (2014) "Temptation, Resolutions, and Regret," *Inquiry* 57: 275-92.

Bratman, M.E. (1987)*, Intention, Plans, and Practical Reason,* Cambridge: Harvard University Press. (Re-issued CSLI Publications 1999.)

Bratman, M.E. (1999a) "Cognitivism About Practical Reason," in M.E. Bratman, *Faces of Intention* New York: Cambridge University Press.

Bratman, M.E. (1999b) "Toxin, Temptation, and the Stability of Intention" in M.E. Bratman, *Faces of Intention* New York: Cambridge University Press.

Bratman, M.E. (2000) "Reflection, Planning, and Temporally Extended Agency," *The Philosophical Review* 109: 35-61.

Bratman, M.E. (2007a) "Valuing and the Will," in M.E. Bratman, *Structures of Agency* New York: Oxford University Press.

Bratman, M.E. (2007b) "Three Theories of Self-Governance" in M.E. Bratman, *Structures of Agency* New York: Oxford University Press.

Bratman, M.E. (2009a) "Intention, Practical Rationality, and Self-Governance," *Ethics* 119: 411-443.

Bratman, M.E. (2009b) "Intention, Belief, Practical, Theoretical," in S. Robertson (ed.) *Spheres of Reason: New Essays on the Philosophy of Normativity* Oxford: Oxford University Press.

Bratman, M.E. (2010) "Agency, Time, and Sociality," *Proceedings and Addresses of the American Philosophical Association* 84:2 7-26.

Bratman, M.E. (2012) "Time, Rationality, and Self-Governance," *Philosophical Issues* 22: 73-88.

Bratman, M.E. (2014) *Shared Agency: A Planning Theory of Acting Together* New York: Oxford University Press.

Bratman, M.E. (2017) "Rational Planning Agency," *Royal Institute of Philosophy Supplement* 80: 25-48.

Bratman, M.E. (2018a) "Introduction: The Planning Framework," in M.E. Bratman, *Planning, Time, and Self-Governance: Essays in Practical Rationality* New York: Oxford University Press.

Bratman, M.E. (2018b) "A Planning Agent's Self-Governance Over Time," in M.E. Bratman, *Planning, Time, and Self-Governance: Essays in Practical Rationality* New York: Oxford University Press.

Broome, J. (2001) "Are Intentions Reasons? And How Should We Cope with Incommensurable Values?' in C. W. Morris and A. Ripstein (eds.) *Practical Rationality and Preference: Essays for David Gauthier* Cambridge: Cambridge University Press.

Broome, J. (2013) *Rationality Through Reasoning* Wiley-Blackwell.

Davidson, D. (2004) "Incoherence and Irrationality," in D. Davidson, *Problems of Rationality* Oxford: Oxford University Press.

Enoch, D. (2006) "Agency, Schmagency: Why Normativity Won't Come from What Is Constitutive of Action," *Philosophical Review* 115: 169-98.

Ferrero, L. (2012) "Diachronic Constraints of Practical Rationality," *Philosophical Issues* 22: 144-164

Frankfurt, H. (1988) "Identification and Wholeheartedness," in H. Frankfurt, *The Importance of What We Care About* Cambridge: Cambridge University Press.

Gibbard, A. (1990) *Wise Choices, Apt Feelings* Cambridge: Harvard University Press.

Goldstein, S. (2016) "A Preface Paradox for Intention," *Philosophers' Imprint* **16**.

Grice, P. (1974) "Method in Philosophical Psychology (From the Banal to the Bizarre)," *Proceedings and Addresses of the American Philosophical Association* 48: 23–53.

Harman, G. (1976). 'Practical Reasoning', *Review of Metaphysics* 29/3: 431–63.

Hart, H.L.A. (2012) *The Concept of Law* 3rd edition Oxford: Oxford University Press.

Holton, R. (2009) *Willing, Wanting, Waiting* Oxford: Oxford University Press.

Kavka, G. (1983) "The Toxin Puzzle," *Analysis* 43: 33-36.

Kolodny, N. (2005) "Why Be Rational?" *Mind* 114: 509-63

Kolodny, N. (2008) "The Myth of Practical Consistency," *European Journal of Philosophy* **16** 366-402.

Kolodny, N. (2018) "Instrumental Reasons," in D. Star, ed., *The Oxford Handbook of Reasons and Normativity* Oxford: Oxford University Press.

Korsgaard, C.M. (2009) *Self-Constitution: Agency, Identity, and Integrity* Oxford: Oxford University Press.

McCann, H. (1991) "Settled Objectives and Rational Constraints," *American Philosophical Quarterly* 28: 25-36.

Morton, J. (2017) "Reasoning Under Scarcity," *Australasian Journal of Philosophy* 95: 543-559.

Morton, J. and Paul S., (2019) "Grit," *Ethics* 129: 175-203.

Nefsky, J. and Tenenbaum, S., (forthcoming) "Extended Agency and the Problem of Diachronic Autonomy".

Núñez, C., (unpublished) "Requirements of Intention in Light of Belief".

Parfit, D. (1984) *Reasons and Persons*   Oxford: Clarendon Press.

Paul, S. (2014) "Diachronic Incontinence is a Problem in Moral Philosophy," *Inquiry* 57: 337-355.

Rabinowicz, W. (1995) "To Have One's Cake and Eat It Too: Sequential Choice and Expected-Utility Violations," *Journal of Philosophy* 92: 586-620.

Raz, J. (2005) "The Myth of Instrumental Rationality," *Journal of Ethics and Social Philosophy*
1:1.

Sartre, J.P. (1975) "Existentialism Is a Humanism," in W. Kaufmann (ed.) *Existentialism from
Dostoevsky to Sartre* rev. and expanded. New York: Meridian/Penguin.

Schroeder, M. (2007) *Slaves of the Passions* Oxford: Oxford University Press.

Setiya, K. (2007) "Cognitivism about Instrumental Reason," *Ethics* 117: 649-73.

Setiya, K. (2014) "Intention, Plans, and Ethical Rationalism" in M. Vargas and G. Yaffe (eds.)
Rational and Social Agency: The Philosophy of Michael Bratman   New York: Oxford
University Press.

Shpall, S. (2016) "The Calendar Paradox," *Philosophical Studies* 173: 801-825.

Simon, H. (1983) *Reason in Human Affairs* Stanford: Stanford University Press.

Smart, J.J.C. (1956) "Extreme and Restricted Utilitarianism," *Philosophical Quarterly* 6: 344-54.

Smith, M (2004) 'The Structure of Orthonomy', in J. Hyman and H. Steward (eds.) *Agency and
Action* Cambridge: Cambridge University Press.

Sobel, J.H. (1994) "Useful Intentions," in J.H. Sobel, *Taking Chances: Essays on Rational
Choice* Cambridge: Cambridge University Press.

Strawson, P. (2003). "Freedom and Resentment," in G. Watson (ed.) *Free Will* Oxford: Oxford
University Press (2nd edn)

Tenenbaum S. and Raffman D. (2012) "Vague Projects and the Puzzle of the Self-Torturer"
Ethics 123: 86-112

Velleman, J.D. (2000) *The Possibility of Practical Reason* Oxford: Oxford University Press.

Velleman, J.D. (2006) "The Centered Self," in J.D. Velleman, *Self to Self* Cambridge:
Cambridge University Press.

Velleman, J.D. (2014) "What Good Is a Will?" in M. Vargas and G. Yaffe (eds.) *Rational and Social Agency: The Philosophy of Michael Bratman* New York: Oxford University Press.

Wallace, R.J. (2001) "Normativity, Commitment, and Instrumental Reason," *Philosophers' Imprint* 1/3: 1–26.

Williams, B. (1981) "Internal and External Reasons" in *Moral Luck* Cambridge: Cambridge University Press.

Yaffe, G. (2000). *Liberty Worth the Name: Locke on Free Agency*. Princeton, NJ: Princeton University Press.

Yaffe, G. (2004) "Trying, Intending and Attempted Crimes," *Philosophical Topics* 32: 505-32.

---

[1] This essay draws from (Bratman 2017). Thanks to Facundo Alonso, Jennifer Morton, Sarah Paul, Kurt Sylvan and Steven Woodworth.

[2] For discussion of the former see (Núñez unpublished); for discussion of the latter see (Yaffe 2004).

[3] This is why it is appropriate to say that these norms articulate rational pressures in the direction of global, on-balance rationality.

[4] This is in the spirit of Grice's strategy of creature construction (Grice 1974). See also (Velleman 2000: chap. 1), (Bratman 2007a).

[5] See the video games example in (Bratman 1987: chap. 8).

[6] I challenge this in (Bratman 1987: 37-39). But I think that even given this strong connection between intention and belief, cognitivism does not work.

[7] For complexities see (Bratman 2009b).

[8] For appeal to forms of inescapability, see (Korsgaard 2009: 68, 82-83), (Velleman 2000).

[9] (Setiya 2014: 74-76) poses this as an issue about "pluralistic rationalism". See also (Enoch 2006).

[10] The latter constitute what I have called self-governing policies. (Bratman 2007b)

[11] (Bratman 2009a). For complexities concerning this inference see (Kolodny 2018).

[12] Concerning the distinction between establishing a reason for conformity to a norm and, further, establishing that it is a norm of rationality, see (Setiya 2014).

[13] Given the hierarchical structure of plans, there can be such interconnections at a higher level despite the absence of such interconnections at a lower level.

[14] We can in this way construct diachronic self-governance out of synchronic self-governance and relevant inter-connections while still maintaining that the self-governance that is of primary interest to us is diachronic.

[15] For alternative approaches see (Ferrero 2012), (Paul 2014).

[16] To have X as an end is, roughly, to have a non-instrumental concern in favor of X. Not all intentions are in favor of such ends, since some intentions are solely instrumental. And, even for a planning agent, not all ends are intentions: it is common to have ends one knows are not co-realizable, whereas having intentions one knows are not co-realizable is a potential rational breakdown.

[17] For appeals to an end of continuity see (Sobel 1994), (Rabinowicz 1995).

[18] A more complete story would also appeal to the agent's expected later regret, if she were now to give into temptation, as a difference between such a temptation case and a version of Kavka's toxin case. (Kavka 1983; Bratman 1999b).

[19] "Complex" in that it is not merely an application of the simple but flawed idea that, quite generally, self-governance of a certain sort must involve the end of self-governance of that sort.

[20] In (Bratman 1987: 24-27) I note that the idea that intentions quite generally provide new reasons threatens to lead to an unacceptable "bootstrapping" of reasons. In (Bratman 2009a; Bratman 2012) I explain how the present proposal is compatible with this point.

[21] (Williams 1981; Schroeder 2007) – though Schroeder appeals to sufficient whereas I appeal to necessary means.

[22] Which is not to say that all self-governance – even self-governance in the pursuit of bad ends -- is good on-balance.

[23] Related ideas are in (Korsgaard 2009), (Velleman 2000).

[24] For further discussion that highlights a parallel with Strawson's (2003) discussion of our framework of reactive attitudes, see (Bratman 2018a). My proposal there is that the end of one's diachronic self-governance (and so of one's synchronic self-governance) plays a role in our reflectively stable planning framework that is to some extent similar to the role, on Strawson's view, that the concern with quality of will plays within our framework of reactive attitudes. While in neither case is the underlying end or concern essential to agency or mind, in each case it plays a fundamental role in our human lives.