

Self-Governance, Means-Ends Coherence and Unalterable Ends*

John Brunero

In "Intention, Practical Rationality, and Self-Governance," Michael Bratman responds to a challenge to the normativity of the requirements of rationality -- a challenge concerning the rational requirement of means-ends coherence and its application to means-ends incoherent agents whose ends and instrumental beliefs are unalterable.¹ In section I of this discussion, I'll explain this challenge and Bratman's response to it, which draws from his recent work on self-governance. In section II, I'll argue against a crucial claim in Bratman's response -- namely, his claim that self-governance is impossible when one's ends and instrumental beliefs are unalterable -- and I'll conclude that his response to this challenge is unsuccessful. In section III, I'll argue that it's doubtful that any theory of self-governance can provide the resources necessary to respond to my objection, and, in section IV, I'll consider how my objection relates to a similar objection to Bratman's view that Bratman himself considers.

I.

Let's first consider what the rational requirement of means-ends coherence requires of you. Suppose you intend to smoke cigarettes, and you believe that a necessary means of doing so is buying a pack, and you believe that you will buy a pack only if you intend to buy a pack, but you don't intend to buy a pack. This combination of attitudes is means-ends incoherent. You could escape this state of incoherence either by coming to intend to buy a pack, or by not intending to smoke, or by revising one of your instrumental

beliefs.² We could formulate the requirement of means-ends coherence as follows, using brackets to indicate the logical scope of "requires":

Means-Ends Coherence (ME): Rationality requires that [if you intend to x , and believe that y -ing is a necessary means to x -ing, and believe that you will y only if you intend to y , then intend to y].

This requirement doesn't require of you that you intend to buy a pack; rather, it requires of you that you *either* not intend to smoke *or* intend to buy a pack *or* revise your instrumental beliefs.

And so, if we were to say that you have a reason to comply with *ME*, it wouldn't follow that you have a reason to intend to buy a pack. Rather, it would merely follow that you have a reason to *either* not intend to smoke *or* intend to buy a pack *or* revise your instrumental beliefs.

But Bratman worries that when your intention to smoke and instrumental beliefs are *unalterable* (in that there is nothing you can do to change those intentions or beliefs) then your reason to comply with *ME* would "transmit" to a reason to intend to buy a pack.³ Bratman endorses the following principle, which would license such transmission:

Transmission-Reasons: If R is a practical reason in favor of X , X is attainable by the agent, and M is a necessary means to or necessary constitutive element of X , then R is a practical reason in favor of M .⁴

In defense of *Transmission-Reasons*, and its applicability to cases of unalterable ends, Bratman writes:

Well, suppose there is an unmovable boulder that prevents me from taking route A to the attainable goal I I have reason to achieve. If, because of this boulder,

taking route B is necessary for that goal, then a reason to achieve that goal transmits to a reason to take route B. And what is plausible, I think, is that the psychological nonmodifiability of intending E makes it appropriate to see that intention as an internal, nonmodifiable analogue of such an unmovable boulder, one that stands in the way of achieving means-ends coherence by dropping that intention. So it is plausible given this psychological nonmodifiability of intending E, a reason for avoiding [Intend E and Not Intend M] transmits to a reason for intending M.⁵

So, if we are to say that you have a reason to either not intend to smoke or intend to buy a pack or revise your instrumental beliefs, but you cannot revise your intention to smoke or your instrumental beliefs, then that reason transmits to a reason to intend to buy a pack.

But Bratman argues that this conclusion is implausible. In the ordinary case in which you can alter your intentions, we don't think that your intention to smoke generates a reason to intend to buy a pack. On Bratman's view, intentions for very bad ends do not generate reasons, nor do intentions that are irrationally akratic.⁶ (I've been using the example of the smoker since that is the original example presented in Setiya's paper. If you don't think smoking is a very bad end, add to the example the fact that it's an irrationally akratic end, or choose a different example substituting in your own favorite very bad end.) Why don't intentions for very bad ends generate reasons? Bratman argues, plausibly, that to treat something as a reason is to see it as "having justifying weight in relevant practical deliberation" and he thinks it's implausible that your intending some very bad end would have any such justifying weight in favor of taking the very bad means to that end.⁷ Why don't intentions for akratic ends generate reasons?

Bratman argues that if they did generate reasons, it would allow for implausible "bootstrapping." Suppose that your best judgment supports alternative A over alternative B, and also supports intending the necessary means to A over intending the necessary means to B, yet you akratically intend to pursue B. If your doing so constituted a reason to intend the necessary means to B, then that reason could have weight sufficient to tip the scales in favor of intending the necessary means to B over intending the necessary means to A. But that would be implausible, especially considering the fact that, by your own lights, alternative A is preferable to alternative B, and the intention that helped tip the scales was itself against your better judgment.⁸

So, in the ordinary case, we don't think that your intention to smoke generates a reason to intend to buy a pack. And, Bratman argues, we should also think the same goes for the case in which your intention to smoke, and your relevant instrumental beliefs, are unalterable.⁹ For this reason, Bratman takes this conclusion -- the conclusion that when your intention and instrumental beliefs are unalterable you have a reason to intend to buy a pack -- to be an implausible one.

If this conclusion is implausible, and *Transmission-Reasons* is sound, then we should reject the claim that you have a reason to comply with *ME*. So, we must reject the idea that requirements of rationality are normative: one could be rationally required to be means-ends coherent and yet have no reason to be means-ends coherent. In this way, the application of the rational requirement of means-ends coherence to agents with unalterable ends and instrumental beliefs poses a challenge to the normativity of rational requirements.

In responding to this challenge, Bratman suggests that we think that the rational requirement of means-ends coherence is normative only if certain background conditions are met. And in the case where you cannot alter your intention to smoke nor your instrumental beliefs, those background conditions are not met, and so there is no reason for you to be means-ends coherent, and so there is no reason that would transmit to a reason to intend to buy a pack. We would thus avoid the implausible conclusion.¹⁰

For this strategy to work, Bratman needs to provide an explanation of, first, why, in normal circumstances, there is a reason to be means-ends coherent, and, second, why, in the special circumstances in which one cannot alter one's end nor one's instrumental beliefs, this reason is not present. Bratman provides such a two-part explanation. First, he argues that we have a reason to govern ourselves, and, in normal circumstances, complying with *ME* would constitute a way of governing ourselves, and so in those circumstances we have a reason to comply with *ME*. Second, he argues that when one's ends and instrumental beliefs *cannot* be altered, this fact "blocks the psychological possibility of relevant self-governance" and so one wouldn't then have this reason (of self-governance) to comply with *ME*.¹¹ Let's consider each part of this explanation in greater detail.

First, let's consider the "normal circumstances" case, and let's assume, with Bratman, that we do have a reason to govern ourselves. Why should we think that compliance with *ME* would constitute a way of governing ourselves? Bratman answers, drawing on some remarks from Harry Frankfurt's "Identification and Wholeheartedness,"¹² that when one fails to comply with *ME*, one fails to *have a place where one stands*. Bratman writes:

Return to a Frankfurtian concern with where I stand. When I recognize inconsistency in my own intentions, I see that in this specific case there is no clear answer to the question, 'Where do I stand?' This question about myself is, with respect to this domain, simply not settled; there is as yet no fact of the matter. We can say something similar about means-ends incoherent plans. If I intend E but I do not now intend known necessary means intending which now I know to be necessary, there is no clear answer to the question, 'Where do I stand?' with respect to E. With respect to this end, there is as yet no relevant fact of the matter about where I stand.¹³

Bratman goes on to argue that having a place where one stands is necessary for self-governance: "it is only if there is a place where you stand that *you* are governing in the corresponding domain...".¹⁴ Since compliance with *ME* is necessary for having a place where one stands, and having a place where one stands is necessary for self-governance, Bratman concludes, "in any particular case, relevant consistency and coherence of intention is a necessary constitutive element in the corresponding self-governance of planning agents like us."¹⁵ And since we have a reason to be self-governing, it follows (by *Transmission-Reasons*) that we have a reason to comply with *ME*.

Second, Bratman argues that when one's intentions cannot be altered there is no psychological possibility of self-governance. Bratman writes:

And indeed, it does seem that if one has an intention that is not susceptible to modification in the light of reflection on reasons and rationality, then that would normally entail that in this specific domain one is not capable of being self-governing. To be self-governing is not only to have a relevant stand but, normally

at least, for one's stand to be psychologically modifiable in the light of relevant reflection.¹⁶

When your intention to smoke is not psychologically modifiable in the light of relevant reflection, you are not capable of self-governance in this particular domain. Hence, it is not the case that you have this reason of self-governance to comply with *ME* in this case. And so it doesn't follow by *Transmission-Reasons* that you have a reason to intend to buy a pack.

In summary, on Bratman's view, when self-governance is possible, there is a reason -- a reason of self-governance -- to comply with *ME*. But when such self-governance is not possible, as when one's ends and instrumental beliefs are unalterable, there is no such reason of self-governance to comply with *ME*.¹⁷

II.

I'll challenge the second part of Bratman's two-part explanation. I'll argue that self-governance *is* possible even when one's ends and instrumental beliefs are unalterable, and so, in these possible cases, there would be a reason of self-governance to comply with *ME*. And this reason would transmit to a reason to intend to buy a pack. So, Bratman does not avoid the implausible conclusion he is aiming to avoid.

Let's start by considering an uncontroversial case of self-governance. Let's suppose Smith, initially, is means-ends incoherent: he intends to smoke, believes that a necessary means to smoking is buying a pack and that he'll buy a pack only if he intends to do so, but doesn't intend to buy a pack. To simplify, let's suppose his instrumental beliefs are unalterable. But Smith is capable of revising his intention to smoke, and

capable of coming to intend to buy a pack. Smith reflects on his current state of means-ends incoherence, considers these two ways of escaping from this state, and deliberates about the pros and cons of each. In light of such deliberation, Smith comes to intend to buy a pack and does so. Now, we may disagree with *how* he is governing himself (since we think he shouldn't have decided to buy a pack) but it is clear *that* he is self-governing. And since he *is* self-governing, self-governance is certainly *possible*. This is a "normal circumstances" case in which Bratman would claim that Smith has a reason of self-governance to comply with *ME*.¹⁸

Now let's vary the case slightly. Jones is exactly like Smith in every way except that the following counterfactual is true of him: were he to decide not to smoke (which, we'll assume, constitutes the formation of an intention not to smoke) someone would intervene and make this intention ineffective in altering his intention to smoke. In other words, the intervener would make it such that Jones's decision not to smoke isn't effective in altering his intention to smoke. Perhaps at the moment of decision, the intervener would administer some potion, or use hypnosis, or re-wire Jones's nervous system to make it such that Jones's decision wouldn't alter his intention to smoke.¹⁹ But Jones knows nothing of this intervention that would occur were he to decide not to smoke. (Indeed, we can assume that his psychology -- his beliefs, desires, intentions, and deliberations -- is *exactly the same* as Smith's: both recognize their state of means-ends incoherence, deliberate about the pros and cons of each alternative, and decide to buy a pack, and so forth.) So, he doesn't decide to buy a pack *because of* this intervention that would occur. Rather, it seems, we should say that, like Smith, he decides to do this *on his own*.

It's important that Jones's psychology is exactly the same as Smith's psychology. It is true that *had* Jones and Smith both decided not to smoke, their psychologies would then be different. Jones's psychology would then be altered by the counterfactual intervener while Smith's would not be. But since Jones and Smith don't *actually* decide not to smoke, there is no actual difference in their psychologies. Both consider the pros and cons of smoking, and in light of their reflection on these reasons, they come to intend to buy a pack. And we can assume that both Jones and Smith are responsive to their assessment of the reasons for smoking: had they seen greater reason not to smoke, they would have decided not to smoke (thereby triggering the intervention in the case of Jones, but not Smith). But, as it happens, they both see greater reason to smoke, and come to intend to buy a pack.

There are two points to note about Jones's case. First, there is only one possible way for Jones to come to comply with *ME*: by coming to intend to buy a pack. Alternate ways of complying are closed off by the counterfactual intervener.²⁰ Since the only way he can comply with *ME* is by coming to intend to buy a pack, if there is a reason to comply with *ME*, that reason would transmit, by *Transmission-Reasons*, to a reason to intend to buy a pack. (Remember that it doesn't matter here that Jones does not *believe* that the only way he can comply with *ME* is by coming to intend to buy a pack; for *Transmission-Reasons*, what matters for the transmission of reasons is that *in fact* the only way he can comply with *ME* is by coming to intend to buy a pack.)

Second, since Smith and Jones do not differ as far as their psychologies go -- they have the same attitudes and deliberate in the same way -- and since Smith's case is a case of self-governance, we should also think that Jones's case is a case of self-governance.

After all, as we said above, both Smith and Jones decide *on their own* to buy a pack. But if Jones *is* governing himself when he comes to intend to buy a pack, then it is certainly *possible* that he governs himself prior to forming this intention. And if self-governance is possible, it follows from Bratman's view that Jones has a reason to comply with *ME*.

But if we now put this second point together with the first -- that Jones has a reason to comply with *ME*, and, that this reason would transmit by *Transmission-Reasons* -- then we get the implausible conclusion that Jones has a reason to intend to buy a pack. So, Bratman doesn't avoid the conclusion he is aiming to avoid.²¹

III.

I've argued that since Smith and Jones do not differ as far as their psychologies go, there is no basis for thinking that Smith, but not Jones, is self-governing. In further support of this claim, I'll argue that neither the pre-theoretical idea of self-governance, nor any leading theories of self-governance, would lend support to the claim that Smith, but not Jones, is self-governing.

Start with the pre-theoretical idea. In his introduction to *Structures of Agency*, Bratman explains the idea of self-governance as follows: "What is self-governance? As an initial, basic step we can say that in self-governance the agent herself directs and governs her practical thought and action. Or anyway, that is the intuitive pretheoretical idea...".²² Bratman goes on to explain that in order to understand what it is for an agent to direct practical thought and action, we need an account of which attitudes have what Bratman calls "agential authority." Attitudes with agential authority are such that when

they guide thought and action, *the agent* directs her thought and action. Such direction is necessary for self-government, but not sufficient. Bratman explains:

For the agent to *govern* her thinking and acting, however, it is not sufficient that she direct them. To govern is to direct in a way that is shaped by what the agent treats as justifying considerations, as reasons. In self-governance, attitudes that have agential authority need to guide relevant thought and action by way of articulating what has, for the agent, justifying significance -- what has subjective normative authority for the agent.²³

So, an account of self-governance would involve, first, an account of agential authority -- that is, an account of which attitudes are such that when they direct thought and action, *the agent* directs thought and action -- and, second, an account of when such direction would be shaped by what the agent treats as justifying considerations.

There are several theories of self-governance.²⁴ On a Platonic theory, very roughly, an agent is self-governing when his thoughts and actions are responsive to his judgments about the good. On a Frankfurtian theory, very roughly, an agent is self-governing when his thoughts and actions are responsive to higher-order attitudes -- specifically a "higher-order volition" that one's first-order desires motivate action. On Bratman's own account, very roughly, an agent is self-governing when his thoughts and actions are responsive to higher-order self-governing policies about which desired ends to treat as reasons in one's motivationally efficacious deliberation.

There's no need for us to go beyond very rough sketches of these theories of self-governance to see that they won't get us what we're looking for -- namely, a basis on which to claim that Smith, but not Jones, is self-governing. These theories all look

toward some psychological features (judgments about the good, higher-order volitions, self-governing policies) in order to understand self-governance. But Smith and Jones, by hypothesis, share the same psychology. And so, on these theories, if Smith is self-governing, then so is Jones. And given that the pre-theoretical idea of self-governance involves the notions of "agential authority" and "subjective normative authority," both of which concern the psychological features of agents, it doesn't seem that *any* theory of self-governance would distinguish between Smith and Jones, who, by hypothesis, share the same psychology.

Additionally, we should note that our problem will not be solved by appealing to Bratman's observation about how self-governance requires that there be a place where the agent stands. Both Smith and Jones transitioned from a state in which there is no determinate answer to the first-personal question, "Where do I stand with respect to smoking?" to a state in which there is. And both could have remained in a state in which there was no answer to this question -- that is, both could have continued to be means-ends incoherent. So, such considerations will not provide a basis for saying that Smith, but not Jones, is self-governing.

Can we provide some explanation of why Bratman is led astray? In the original example from Kieran Setiya, the smoker suffers from some psychological incapacity -- a compulsion or addiction to cigarettes -- that seems to block *both* the possibility of self-governance *and* the possibility of his revising his intention to smoke. (While Setiya doesn't explicitly say that this is a case of compulsion or addiction, Bratman clearly reads Setiya's example this way.²⁵) Since there is a compulsion or addiction in play, it seems reasonable to think that the possibility of self-governance is blocked (and the various

theories of self-governance sketched above would support the claim that such compulsions or addictions can, and usually do, block the possibility of self-governance in a certain domain²⁶). But it *need not* be the case that what blocks the possibility of one's revising an intention also blocks the possibility of self-governance. In Jones's case, what blocks the possibility of his revising his intention to smoke (namely, the intervention that would occur were he to decide not to smoke) *doesn't also* block the possibility of self-governance. And that, I've argued, is troublesome for Bratman's proposal for avoiding the implausible conclusion he is aiming to avoid.

IV.

In a footnote, Bratman hints at the possibility of an objection along the lines I've sketched here but declines to pursue the issue.²⁷ That footnote occurs in the midst of a discussion of a similar objection -- one which aims to show that an intention could be unalterable without blocking the possibility of self-governance. Bratman notes that if an end is a Frankfurtian 'volitional necessity' it is unalterable but doesn't block, and indeed might play a crucial role in, self-governance. I think that consideration of volitional necessities complicates issues unnecessarily, and that the objection I've developed in this discussion avoids these complications and thereby amounts to a more formidable objection to Bratman's proposal than the objection he considers.

What are the complications? First, it's not clear that *Transmission-Reasons*, which is plausible on our ordinary understanding of necessity, would be plausible for volitional necessities. A person with a volitional necessity, according to Frankfurt, "accedes to [the constraining force] because he is *unwilling* to oppose it, and because,

furthermore, his unwillingness is *itself* something which he is unwilling to alter."²⁸ Now suppose that both route A and route B get you to where you have conclusive reason to go, but A will save you a substantial sum of money. If you now tell me that there's an immovable boulder blocking A, I'll agree that you have conclusive reason to take B; your reason here transmits. But if you tell me you are unwilling to alter your unwillingness to take A, I might not agree that you have conclusive reason to take B; I might think you should be willing to alter your unwillingness instead. Now perhaps *Transmissions-Reasons* applies here and perhaps it doesn't. A virtue of the objection I've developed in this discussion is that it doesn't rely on a sense of necessity having to do with an agent's unwillingness, and so the applicability of transmission principles isn't called into question. For that reason, I think it presents a more formidable challenge to Bratman's view than the volitional necessities objection he considers.

Second, some philosophers think that our reasons depend in important ways on features of our psychology. They might think it implausible that a mere intention could bring about a reason to intend the means, but not think it implausible that an intention that is volitionally necessary, or otherwise reinforced by other elements of one's psychology, could bring about such a reason. The objection I've developed avoids this complication too: Smith and Jones, by hypothesis, share the same psychology. What makes Jones's end unalterable isn't his unwillingness to revise that end, nor is it any other feature of his psychology.

Third, a volitionally necessary intention, as defined by Frankfurt, cannot be an akratic intention.²⁹ But intentions that are unalterable because of the presence of a counterfactual intervener can still be akratic intentions. (We could, for instance, suppose

that both Smith and Jones are akratic in having their intentions to smoke.) This is important because one reason that Bratman rejects the idea that intentions generate reasons in general is that he thinks that akratic intentions do not generate reasons, regardless of whether those intentions are alterable or unalterable.³⁰ But if we assume that Jones is akratic in his intention to smoke, and that Jones's intention to smoke is unalterable because of the counterfactual intervener, but that (as I've argued) self-governance is possible for Jones, then, on Bratman's view, Jones's intention to smoke would indeed, implausibly, generate a reason for Jones to intend to buy a pack. This is a serious challenge for Bratman's view. No similar challenge to Bratman's view can come from consideration of volitional necessities. For we cannot, on Frankfurt's definition of 'volitional necessity', assume *both* that Jones is akratic in his intention to smoke *and* that Jones's intention to smoke is unalterable because it is volitionally necessary.

In conclusion, the objection I've developed in this discussion avoids the complications that come along with consideration of volitional necessities. The objection shows that what blocks the possibility of altering an end may not block the possibility of self-governance, and this presents a difficulty for Bratman's proposed solution to the problem for the normativity of means-ends coherence posed by its application to unalterable ends.

Notes

* An earlier version of this paper was presented at the 2010 Central Division meeting of the American Philosophical Association, where I received helpful comments from Sergio Tenenbaum, and at the 2009 Central States Philosophical Association meeting, where I received helpful comments from Kenneth Williford. I also owe much to comments and advice I received from Dana Tulodziecki, and from two anonymous reviewers for *Ethics*, and its Editor, Henry Richardson.

¹ Michael Bratman, "Intention, Practical Rationality and Self-Governance" *Ethics* 119 (2009): 411-443. The problem Bratman is responding to is a slight variation on a challenge to the normativity of rational requirements originally introduced by Kieran Setiya, "Cognitivism about Instrumental Reason" *Ethics* 117 (2007): 646-673.

² Though, if you revise either of your instrumental beliefs -- that is, here, your belief that a necessary means to smoking is buying a pack, or your belief that you will buy a pack only if you intend to buy a pack -- then you would violate some other requirement of *theoretical* rationality, assuming you take these beliefs to be well-supported by the available evidence.

³ In Setiya's original presentation of this worry, the agent with an unalterable end is such that there is "nothing [he] can do to change [his] intention to smoke." In particular, there is no decision he could make that would affect his intention. For example, we are to "suppose that [his] intention to smoke is sufficiently robust that even if [he] decided not to smoke, the resulting conflict of intentions would be resolved in its favor: [he] would still intend to smoke." See Setiya, "Cognitivism about Instrumental Reason," 661.

⁴ Bratman, "Intention, Practical Rationality and Self-Governance," 424. For similar transmission principles, see Setiya, "Cognitivism about Instrumental Reason," 656, 660, and Patricia Greenspan, "Conditional Oughts and Hypothetical Imperatives" *Journal of Philosophy* 72 (1975): 259-76, 265.

⁵ Bratman, "Intention, Practical Rationality and Self-Governance," 425.

⁶ *Ibid.*, 415-416.

⁷ *Ibid.*, 416.

⁸ *Ibid.*, 415-416. For additional, related arguments concerning the relation between intentions and reasons, see Michael Bratman *Intentions, Plans and Practical Reason* (Cambridge, MA: Harvard University Press, 1987), 23-27.

⁹ Bratman, "Intention, Practical Rationality and Self-Governance," 426.

¹⁰ *Ibid.*, 428-435.

¹¹ *Ibid.*, 434.

¹² Reprinted in Harry Frankfurt, *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988): 159-196, 166.

¹³ Bratman, "Intention, Practical Rationality and Self-Governance," 431.

¹⁴ *Ibid.*, 431.

¹⁵ *Ibid.*, 431. Bratman here claims that consistency and coherence of intention are necessary constitutive elements of self-governance. His argument doesn't require him to say (and he doesn't say) that self-governance is *itself* a further end that agents must adopt; rather, his argument is simply that the having of consistent and coherent intentions is a necessary constitutive element of self-governance, so, since we have a reason to be self-governing, we have a reason to have consistent and coherent intentions.

¹⁶ Ibid., 434-435. As Bratman notes in n. 64, the "normally at least" here is put in place to accommodate cases in which one could be self-governing by responding to the fact that one's end is unalterable, perhaps by locking up the cigarette cabinet, in the case of the smoker.

¹⁷ As Bratman acknowledges, his view does concede that there are some cases in which *ME* is applicable although there is no reason to comply with *ME*; as he puts it, we are "agreeing with the myth theorist [that is, the theorist who thinks the idea that rationality is normative is a myth] since we are granting that there are cases to which Means-Ends Coherence applies though there is not the cited reason for conformity to Means-Ends Coherence." See Bratman, "Intention, Practical Rationality and Self-Governance," 433.

¹⁸ That reason wouldn't transmit, by *Transmission-Reasons*, to a reason to intend to buy a pack since intending to buy a pack is not necessary for complying with the reason he has to be means-end coherent: he could instead not intend to smoke.

¹⁹ These three methods of intervention are mentioned by Harry Frankfurt in his "Alternative Possibilities and Moral Responsibility" *Journal of Philosophy* 66 (1969): 829-839, 835-836 -- a paper that is the obvious inspiration for the line of argument advanced here. Though my argument draws from Frankfurt's example, we should note an important difference between Frankfurt's counterfactual intervener and my own. In Frankfurt's example (changing the names), one individual, Black, wants to get another, Green, to perform some action, and he is an excellent predictor of whether or not Green will perform that action on his own. If Black comes to believe that Green will not perform the action on his own, then he'll intervene and take the necessary steps to make sure Green performs that action. But since Green decides on his own to perform that

action, and does perform it, Black never has to intervene, and Green never becomes aware of the fact that Black would have intervened. Frankfurt's point is that even though Green could not have done otherwise (since if he hadn't decided to perform that action, Black would have taken the steps to ensure he did perform it) he is nonetheless morally responsible for his action. It would be inappropriate to excuse him from blame (or deny him praise) for his action. In Frankfurt's example, Black excludes *all but one* possibility for Green. That's why Frankfurt's example is a challenge to the Principle of Alternate Possibilities, according to which one is morally responsible only if one could have done otherwise. In our example, however, the counterfactual intervener doesn't exclude all but one possibility: Jones could intend to buy a pack and he could continue to be means-ends incoherent; alternate possibilities remain open. But, importantly, the counterfactual intervener ensures that there is only one possible way for Jones to *comply with ME*: intending to buy a pack.

²⁰ We'll assume that *whatever* means he would take to removing the intention to smoke would be interfered with. One such obvious means, considered above, is deciding not to smoke. But perhaps there are others.

²¹ In my example, intending to buy a pack is, for Jones, a necessary means of complying with *ME*, but Jones *doesn't know* that it is a necessary means since he doesn't know of the counterfactual intervener, and so when Jones deliberates and comes to intend to buy a pack, it is clear that, like Smith, he is self-governing. Now perhaps one could construct other examples -- perhaps examples which don't involve counterfactual interveners -- in which intending to buy a pack is a necessary means for complying with *ME*, but the agent doesn't know this and instead believes there are alternate possible ways of complying

with *ME*. And perhaps in these cases, when the agent comes to intend to buy a pack, it's clear that he is self-governing. For example, perhaps Brown is exactly the same as Smith except in that he has a psychological compulsion that is *unknown to him* and renders the revision of his intention to smoke impossible. But Brown mistakenly believes himself capable of revising his intention to smoke and capable of coming to intend to buy a pack. Brown then proceeds to deliberate about the pros and cons of each of these supposedly possible ways of proceeding in the same way that Smith and Jones do, and decides to buy a pack and does so. Intuitively, Brown is self-governing, even though he cannot alter his intentions nor his instrumental beliefs.

Thus, it seems that one could construct other examples without counterfactual interveners to challenge Bratman's proposal. However, I think the counterfactual intervener case makes the point quite clearly, since there is *no* psychological difference between Jones and Smith and hence nothing that could ground the claim that Smith, but not Jones, is self-governing. There is a psychological difference between Brown and Smith since Smith does not have Brown's unknown compulsion, though I suspect (but will not argue for it here) that this difference wouldn't be significant enough to ground the claim that Smith, but not Brown, is self-governing.

²² See Michael Bratman, "Introduction" in *Structures of Agency* (Oxford: Oxford University Press, 2007), 4. Though self-governance is a central theme throughout *Structures of Agency*, it figures most centrally in "Planning Agency, Autonomous Agency" and its companion piece "Three Theories of Self-Governance."

²³ Bratman, "Introduction" in *Structures of Agency*, 4-5. See also Michael Bratman, "Rational Intention" *Philosophical Explorations* 12 (2009): 227-241, sec. 3.

²⁴ See especially Bratman's "Reflection, Planning, Temporally Extended Agency," "A Desire of One's Own," and "Three Theories of Self-Governance" all in *Structures of Agency*, for elaboration on and criticism of these views.

²⁵ Bratman, "Intention, Practical Rationality and Self-Governance," 423, 439.

²⁶ Though perhaps not, if that compulsion or addiction is unknown to the agent and doesn't influence the agent's deliberations, as in my example of Brown in n. 21 above.

²⁷ Bratman, "Intention, Practical Rationality and Self-Governance," 440, n. 70.

²⁸ See Harry Frankfurt, "The Importance of What We Care About", in *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988), 87.

²⁹ Bratman acknowledges this. See "Intention, Practical Rationality and Self-Governance," 441.

³⁰ Bratman, "Intention, Practical Rationality and Self-Governance," 415, 426.