Intention rationality

Michael E. Bratman*

School of Humanities and Sciences, Philosophy Department, Stanford University, Stanford, California, USA

The practical thought of planning agents is subject to distinctive rationality norms. In particular, there are norms of intention consistency and of means–end coherence. I discuss the normative significance of these norms and their relation to practical reasons. I seek a path between views that see these norms as, at bottom, norms of theoretical rationality, and views that see the idea that these norms have distinctive normative significance as a ‘myth’. And I seek to distinguish these norms from principles about the transmission of practical reasons. In the end, my view draws on claims about what is involved in being a self-governing planning agent.

**Keywords:** intention; planning; practical rationality; practical reasons; self-governance

1. Intentions are fundamental elements in a psychology of planning agency. These intentions are guided by the agent’s (at least, implicit) acceptance of distinctive norms, thereby tend to conform to those norms, and are rationally criticizable when they violate these norms. Central among these norms are norms of intention consistency and means–end coherence. The former enjoins consistency among one’s intentions given one’s beliefs; the latter requires that one’s structure of intentions be filled in with specifications of means and the like as one sees this to be necessary in the pursuit of intended ends. Both norms are *pro tanto* demands: their violation involves a rational failure, but there can be cases in which other considerations nevertheless take precedence. In being guided by the (implicit) acceptance of these norms, a planning agent’s psychic economy tracks conformity to these norms and tends to issue in such conformity.¹

These are norms of synchronic rationality of planning agency. There is also, I think, a norm of rational stability of intentions and plans over time. Here, however, my focus will be on these norms of synchronic rationality. I leave for another occasion a discussion of how the ideas to be developed in this essay might apply to a norm of cross-temporal stability.

A complication is that, as we have learned from Donald Davidson (2004), a certain degree of conformity to such norms of consistency and coherence is a condition for being interpretable as a person with beliefs and intentions. Cases of clear and explicit violations of these norms will be unusual, and given an apparent case of such a violation there will be epistemic pressure to re-think our understanding of the agent’s psychological profile. Nevertheless, I take it that these are norms whose violation remains a possible form of irrationality.

---

¹Email: bratman@stanford.edu

ISSN 1386-9795 print/ISSN 1741-5918 online
© 2009 Taylor & Francis
DOI: 10.1080/13869790903067717
http://www.informaworld.com
Further, I take it that we – adult human agents in a broadly modern world – are such planning agents, and our practical thought and action is normally guided by these norms. A threat of means–end coherence poses deliberative problems; and a threat of inconsistency induces a filter on options to be considered in deliberation. Being guided in these ways by our (implicit) acceptance of these norms is an aspect of the proper functioning of planning in our agency.

But why should we, on reflection, care about conforming to these norms? Why is this important to us?

Well, being guided by these norms is, as noted, an essential element of our planning agency, and there is much to be said in favor of being a planning agent. Planning structures are universal means: they help promote a wide range of complex and temporally extended ends (Bratman 1987). Planning structures help to support and to constitute forms of unity of agency that are central to our cross-temporal self-governance – a point to which I will return (Bratman Forthcoming a). And planning structures are central elements in forms of sociability we care about. (Bratman 2009b, Forthcoming c). In these ways, planning structures help to support and to constitute complex and effective forms of agency that we highly value. And part of being a planning agent is being guided by our (implicit) acceptance of the cited norms.

These reflections concern the impact and significance of general modes of thought and action characteristic of planning agency. And this is an important element in our understanding of the significance of the cited norms of intention rationality. Still, this is not a complete answer to our question. In particular, we need also to go on to ask why we should care in the particular case about conforming to these norms in that case. And our reflections so far on general modes of thought and action do not yet fully answer this question. After all, one thing we have learned from debates about rule utilitarianism and worries about ‘rule worship’ is that the step from support for a general mode of thinking to support for a specific instance is fraught (Smart 1967). This is because it remains possible that the features appealed to as support for the general mode of thinking might not be present in a specific, particular case.

So: why should we care in the particular case about conformity to these norms? Well, in the particular case we will have certain intended ends, and we will normally think we have reasons for those ends. Further, we will see that, normally, consistency and coherence of relevant intentions and plans is instrumentally needed successfully to pursue those intended ends. So we will normally thereby be in a position to conclude that we have a reason – an instrumental reason – for such consistency and coherence of intention.

True enough. But this is still a limited answer. First, it seems possible to intend some end without thinking one has a normative reason for that end. This is a lesson we learn from reflection on cases of weakness of will and related cases of intentions motivated by certain kinds of emotions or compulsions. Perhaps – to use an example along lines suggested by David Hills – in a moment of pique I decide to ignore you at the party. Later I come to think that there is really no reason to do this; yet out of stubbornness I stick with my intention to ignore you, while knowing all too well that I have no justification for doing this. Nor do I think that simply by intending to ignore you I thereby have a reason for doing this that I did not otherwise have, that my intention bootstraps a new normative reason for this action (Bratman 1981). Still, I do intend to ignore you, and this intention is poised to guide my thought and action. Such an intention engages norms of consistency and coherence, even though I am not in a position to judge that there is the cited kind of instrumental reason for conformity to those norms in this particular case. So we can still wonder whether – and, if so, why – I should care about conformity to these norms of consistency and coherence as they apply to this intention in this particular case.
Second, this is, in any case, to see consistency and/or coherence of intention in the particular case as instrumentally useful with respect to supposed reasons for the specific individual ends in question. But even if this were true, in a wide range of cases we would still be left wondering whether or not consistency and coherence of intention in the particular case have a normative significance that goes beyond this instrumental impact. We would still wonder if consistency and coherence have, in the particular case, a non-instrumental normative significance that is distinctive in the sense that it does not depend on their promotion of these supposed, specific reasons for the specific individual ends. This question remains even after we have noted the significant virtues of general modes of thinking associated with planning agency, and even after we have noted that, in a range of particular cases, consistency and coherence of intention are instrumentally supportive of what we have reasons to do.

One response at this point is to claim that consistency and coherence of intention do not have their own distinctive, non-instrumental normative significance: the thought that they do is a ‘myth’. A view along these lines has recently been defended both by Joseph Raz and by Niko Kolodny. Such a view can grant that there is much to be said in favor of associated, general habits of mind. And it can grant that on most occasions the agent will sensibly suppose that there are reasons for her specific intended ends that transmit to instrumental reasons for associated consistency and coherence of intention. But such a view insists that it is a myth that consistency and/or coherence of intention in the particular case have their own distinctive, non-instrumental normative significance.

Such a myth theory is in tension with our commonsense understanding of our own practical rationality. When we think of ourselves as inconsistent or incoherent in our intentions in a particular case, we normally think of ourselves as failing to satisfy, in that very case, basic demands of reason. We think this whether or not conformity to those demands would have been instrumentally useful in the pursuit of other things. Indeed, such thoughts seem part of the way in which the normal consistency-and-coherence-tracking functioning of our intentions and plans is stable under reflection. Perhaps, in the end, we could be convinced that such thoughts express unsupported myths. But this would be a high philosophical price to pay, and before we consider paying it we need to reflect carefully on the philosophical alternatives.

One alternative that has loomed large in the recent literature is to see these norms of intention rationality as, at bottom, norms of theoretical rationality on associated beliefs. Intentions are to be consistent because intention involves belief, and beliefs are to be consistent. Intentions are to be coherent because intention involves belief, and beliefs are to be coherent. This is cognitivism about these norms of intention rationality. Cognitivism needs a tight connection between intention and belief. Further, if cognitivism is to account in a straightforward way for, in particular, the consistency demand on intention, it seems that what it will need is a tight connection between intending A and believing, in particular, A. This is because a connection to a weaker belief – e.g. a belief only that it is possible that one will A – will have trouble accounting for the demand that one not both intend A and intend B when one believes A and B are not co-possible. After all, it is normally fine to believe that A is possible and believe that B is possible while also believing that it is not possible that A and B. However, the idea that intending to A involves the belief that one will A is a rather strong assumption, one that has its own philosophical costs. After all, it seems that one can, for example, intend to do certain things later while retaining some doubt about whether one will remember when the time comes.

But let us put this concern to one side and ask whether, on the assumption of such a strong connection between intention and belief, cognitivism would be an attractive strategy for avoiding a myth theory about these norms of intention rationality.
Well, such cognitivism would be a reasonable response to a myth theory only if it could establish that consistency and/or coherence of belief in the particular case has its own distinctive normative significance. If you were also a myth theorist about the distinctive significance of these norms of theoretical rationality, such cognitivism would not block a myth theory about the cited norms of practical rationality. And, indeed, Raz (2005, section 3) and Kolodny (2008b) are also myth theorists about norms of consistency and coherence of belief.

Still, it may seem that it is easier to explain the distinctive normative force of demands of consistency and coherence on belief than it is to explain analogous demands on intention. Belief, we can try to say, aims at truth: an attitude that was not located within a psychic economy that was broadly truth-tracking would not be belief. (Fantasy, in contrast, need not be truth-tracking.) Further – though it is difficult to nail this down – this aspect of the basic functioning of belief seems somehow to support the idea that there is a norm of correctness on belief, a norm that says that a belief is correct if and only if it is true. And consistency and coherence have a close relation to truth. So perhaps we can see appeal to the truth-aim of belief as a key to understanding the distinctive normative pressure for consistency and coherence of belief in the particular case.

However, even if this is a way of avoiding a myth theory about these norms of theoretical rationality, there remains a further problem with the effort to appeal to these theoretical norms to explain the distinctive normative force of demands of consistency and coherence on intention. This problem is posed by the apparent possibility of false beliefs about one’s own intentions. In particular, it seems that I can sometimes believe I intend a certain means and yet not in fact so intend. But then it will be possible for my relevant beliefs (where these include my false belief that I intend the means) to be fully coherent even though my intentions do not satisfy the demand for means–end coherence since I do not in fact intend needed means (Bratman 1981, note 4).

There are several efforts in the literature to defend cognitivism from this objection, and I have discussed these efforts elsewhere (Bratman 2009a; Forthcoming a). Here I will simply take it for granted that, as I think, this objection at the least poses a serious problem for such cognitivism. Now, one motivation for trying to defend cognitivism in the face of this objection is that one supposes that the alternative to cognitivism is a myth theory. Indeed, as I will be noting along the way, this seems to be the view, for somewhat different reasons, of R. Jay Wallace (2001) and Kieran Setiya (2007). And one of my main aims in this essay is to challenge this supposition by defending, in broad outlines, an alternative strategy for avoiding a myth theory. This alternative strategy is to see these norms of intention rationality as distinctive practical norms, rather than as norms that derive from associated norms of theoretical rationality.

But why think there are these practical norms of intention rationality conformity to which in the particular case has its own distinctive, non-instrumental normative significance? We do not think that there are such norms on ordinary desire. It can sometimes be straightforwardly rational to desire A and desire B while knowing that A and B are not co-possible. Just think about your desire for sweets and your desire for a slim figure. Of course, such conflicts – which pervade our lives – call for decision in the particular case. But desires are not decisions. Again, even if I desire some end, I may be under no rational pressure to give any thought at all to means; and, indeed, I might know that any realistic means to what I desire would itself be something I do not desire. Such is human life, replete with unrealistic yearnings. Granted, once I decide to pursue the end things change in important ways. But, again, desire is not decision.

What we need, then, is to understand why a transition, normally by way of decision, from desire to intention newly engages distinctive rational norms of consistency and
means–end coherence. The cognitivist tries to provide an answer to this: the transition from desiring A to intending A ensures a belief that A, and that belief engages theoretical norms of consistency and coherence. But we have seen that this strategy is problematic; and we are currently trying to develop an alternative strategy, one that sees the relevant rationality norms as distinctive practical norms.

Here it is important to note that it will not suffice to appeal to some sort of principle of transmission of reasons for ends to reasons for what is necessary for those ends: such a transmission-of-reasons principle would still not get us the cited norms of coherence and consistency of intention. After all, as we noted earlier, it seems that one can intend an end without believing there is a reason for that end, and so without believing there is a reason to transmit. But such intentions still engage the cited norms of consistency and coherence.

This point – that intending can come apart from believed reasons – is emphasized by R. Jay Wallace (2001, 3–10; 15–16) in criticizing what he takes to be Christine Korsgaard’s (1997) strategy of focusing on a transmission-of-reasons principle. And I am agreeing with Wallace that, for this reason, such a focus would not suffice as a defense of the cited norms of consistency and coherence against the myth theorist. Wallace, however, seems to infer from this that the further step we need in order to avoid a myth theory is a step to cognitivism. And this is an inference I mean to be challenging.

Now, when we considered the theoretical demands of consistency and coherence on belief, we noted the possibility of an appeal to the idea that belief aims at truth. Might we pursue an analogous strategy here, and appeal to the aim of intention?

Well, how would we determine the aim of intention? The answer, I take it, is that we look at the characteristic roles of intention in our psychic economy. On the planning theory, these roles are the coordinated and effective control of temporally extended action. Further, this coordinated effective control proceeds by way of the incremental specification and articulation of the system of intentions – for example, by specifying needed means and avoiding specifications that are in conflict. An attitude that was not guided by a system that tracked and supported such coordinated effective control, by way of such incremental specification, would not be an intention. (In contrast, ordinary desire need not track or support systematic coordination.) So we might try to say that whereas belief aims at truth, intention aims at coordinated, effective control of temporally extended action by way of the incremental specification of the system of intentions. And, assuming that the agent has relevant true beliefs, achieving this aim of intention will involve a system of intentions that is coherent and consistent. So perhaps we can say that the norms of intention consistency and coherence are supported by the aim of intention in a way that parallels the support for belief consistency and coherence that is grounded in the truth aim of belief.

Now, I do think some such appeal to the aim of intention – an appeal that is to some extent parallel to an appeal to the aim of belief – is plausible (Bratman 2009a; Forthcoming a). But what we need to know for present purposes is whether such an appeal to the aim of intention gives us something more to say to the myth theorist.

Suppose I am wondering why I should care right now about the consistency and coherence of my intentions on the present, particular occasion. My intentions are attitudes that are embedded in a distinctive functional economy and thereby – or so we are supposing – can be said to aim at coordinated effective control by way of incremental specification. And let us suppose that intention consistency and coherence are indeed central to such coordinated effective control by way of incremental specification. Nevertheless, it seems that I might know all this and still wonder why I should care whether my current thought is on this occasion the expression of, in particular, intentions that have this aim, rather than of
different practical attitudes that do not have this aim. Why not revert to non-planning agency? And so far the answer we have provided to this question is an answer at the level of general modes of thinking: there are many good features of being, in particular, a planning agent. While this is of the first importance, it does not get us further than where we were before: namely, that there is good reason for the general patterns of thought and action characteristic of planning agency. So appeal to the aim of intention does not yet get us to an account of the distinctive, non-instrumental normative force of the norms of intention consistency and coherence in the particular case, and so does not yet provide an answer to the myth theorist.

Does the same worry arise for a cognitivist who appeals to the aim of belief to support theoretical norms of consistency and coherence of belief in the particular case? Well, it might perhaps be possible for such a cognitivist to argue that some truth-tracking attitude is essential to cognition quite generally. So perhaps it can be argued that it is not a real option for a cognizer to abandon truth-tracking attitudes and retreat entirely to non-belief cognitions that do not aim at truth. Fantasy as a complete cognitive strategy is not an option. In contrast, it is not in general necessary that an agent be a planning agent; indeed, it seems that many non-human agents are not planning agents but, rather, only desire-belief agents. So a possible retreat to non-planning agency may seem a more pressing possibility with which we need to come to terms (which is not to say that we have it in our power simply to decide to be non-planning agents).

In any case, whatever the final verdict about such appeals to the truth-aim of belief (a matter I do not try to settle here), when we turn to intention and seek an account of the norms of intention rationality as practical norms, appeal to the aim of intention, while of significance, does not yet give us an account of the distinctive, non-instrumental normative force of these norms in the particular case.

Can we simply say that the distinctive, non-instrumental normative force of these norms in the particular case is as fundamental as things get, and it is a mistake to demand a further explanation? This would be a kind of quietism about this normative force. And the problem is that what we have seen is that there is a lot we can say in support of these rationality norms and of related (though different) ideas, without saying that these rationality norms have distinctive, non-instrumental normative force in the particular case. In particular, we can say:

1. These rationality norms are aspects of a general mode of functioning of planning agency for which there are powerful reasons; and the intentions that are elements in this functioning can thereby be said to aim at coordinated effective control of temporally extended action by way of incremental specification;
2. Conformity to these rationality norms is normally conducive to that for which we independently have reasons;
3. There are relevant principles of the transmission of reasons from ends to means.

The myth theorist says that (1)–(3) are all that there is to say here. The quietist insists they are not. But in the absence of a further, substantive story, this quietist response may seem to be just a kind of stubbornness.

We would have a substantive response if we could argue that there are, systematically, distinctive, non-instrumental practical reasons to conform to these norms in the particular case, where these reasons are not just reasons for the associated general habits of mind, or reasons tied to theoretical demands on associated beliefs or derived from the usefulness of consistency and coherence in the particular case. As we might say: there is,
systematically, a substantive, distinctive, non-instrumental practical consideration at stake in the particular case when the issues of intention consistency and coherence arise in shaping a planning agent's deliberative agenda. This would not be to reduce these rationality norms simply to a claim about reasons for conformity; but it would be to cite an important link between practical rationality and practical reasons. And our question is whether we can defend such a link between intention rationality and reasons for conformity in the particular case.

2.

I think we can defend some such link; but both the link, and our route to it, will be complicated. To lay the groundwork, there are three preliminary matters we need to address.

The first involves reflecting on the very idea of a normative, practical reason. What is this idea, anyway? Well, I take it that the relevant idea coheres with T.M. Scanlon’s comment that ‘a reason for something ... [is] a consideration that counts in favor of it’ (Scanlon 1998, 17). A normative practical reason is a consideration that counts in favor of intention or action. Further, I think we should say that to judge that R is a normative reason for A is in some sense to endorse R as ‘a consideration that counts in favor of’ A.9

This leaves it open how to understand the nature of this endorsement. On an expressivist view along lines developed by Allan Gibbard, this endorsement involves the ‘acceptance of norms that say to treat R as weighing in favor of’ A (Gibbard 1990, 163). On a realist view, this endorsement involves the cognitive recognition of a normative fact. While I myself find a broadly expressivist approach more plausible, my aim here is to articulate a link between intention rationality and reasons, a link that is available to different views about such endorsement.

The next preliminary involves two claims in our normative philosophy about the relation between intentions and reasons. The first is that sometimes one’s intentions, plans and projects provide reasons for action that one would not otherwise have had. Going to law school and going to a philosophy PhD program are both valuable things to do; but it is because you are embarked on, in particular, a career in philosophy that you have reasons for actions in pursuit of a philosophy PhD, reasons that you do not have for a course of action in pursuit of a JD. Nevertheless – and this is the second claim – not all intentions provide such reasons for action. As I have already suggested, the agent herself might sensibly suppose that her weak-willed intention, or her intention motivated by some sort of compulsion or powerful emotion, provide no reason – no normative, justifying reason – for action. Further, once we are given the idea that to judge there is a reason is in a characteristic way to endorse, we are in a position to say that someone’s intention in favor of something very bad may fail to provide a normative reason for his pursuing that bad thing. So while intentions sometimes provide normative reasons for action, they do not always provide such reasons: they do not in general bootstrap new justifying reasons for what is intended or for means to that.10

Return now to the norms of consistency and coherence of intention. We want to know if – and if so, how – we can provide a distinctively practical interpretation of these norms, one that steers a path between cognitivism and a myth theory, and one that does not simply appeal to a transmission-of-reasons principle. Following John Broome, we can say that these norms have ‘wide scope’: they concern, in the basic case, packages of attitudes (Broome 2000). The consistency norm tells us not to have a package of attitudes that includes an intention to A, an intention to B, and the belief that it is not possible to do both. The coherence norms tells us, roughly, not to have a package of attitudes that
includes intending E, believing that M is a necessary means to E, and yet still not intending M. The cognitivist tries to see packages that violate these norms as ineluctably involving packages of beliefs that violate corresponding theoretical norms. In contrast, we want to know if we can say that we have a distinctive, non-instrumental practical reason to conform to these wide-scope demands of intention coherence and consistency in the particular case. If we could say this, we would go beyond appeal to reasons in favor of general modes of practical thinking, without retreating to an appeal to theoretical reasons for coherence and consistency of belief. And my third preliminary point is that in trying to say this we need to be sensitive to a problem about the transmission of reasons.

Let’s focus initially on the norm of means–end coherence. And, though we will have reason below to express this idea in a more qualified way, let’s consider:

(R) There is a distinctive, non-instrumental practical reason in favor of the following:
if you intend E and believe M is a necessary means to E then you intend M

where this reason does not require that there is a reason for intending E: (R) is not a transmission-of-reasons principle. And let us suppose that in the particular case:

(1) You intend E and believe M is a necessary means to E.

So one way in which you can conform to the practical reason cited in (R) is to intend M.
This poses the question of whether it follows from (R) and (1) that:

(2) You have a practical reason to intend M.

And if (2) were to follow from (R) and (1), we would have a problem. After all, this route from intending E to a reason to intend M would be quite general, requiring only the belief that M is a necessary means to E. So if we were to say that in such cases, quite generally, you have a reason to intend M, we would be saying that intending E quite generally provides a reason for intending believed necessary means. But this is in tension with our earlier claim that intentions do not in general provide such reasons.

However, we have learned from John Broome that it is quite plausible that an inference from (R) and (1) to (2) – what Broome calls a ‘factual detachment of a reason’ – is not a valid inference (Broome 2005, 5). After all, it is consistent with (R) and (1) that it remains up to you whether to continue to intend E. So it is plausible to say that, though M is believed to be necessary for E, intending M is not in the relevant sense necessary for conforming to the reason cited in (R), since there is available to you an alternative way to conform to that reason: namely, giving up intending E.

To get deductively from (R) to (2) by way of an acceptable transmission of reasons principle, we need a further premise that establishes that intending M is appropriately necessary for conforming to the reason in (R). And what we learn from Broome is that it is plausible that the relevant kind of necessity of intending M is not ensured simply by the factual presence of the intention in favor of E, given that it remains up to you whether to continue to intend E. So far, then, we can protect the claim that intentions do not in general provide normative practical reasons for intending the means.

Nevertheless, a problem looms. Suppose you intend end E but that because of a psychological incapacity you do not have the ability to change this intention. In an example drawn from Keiran Setiya, suppose you have a compulsive desire to smoke, that this induces an intention to smoke, and that the compulsiveness of your desire undermines your normal capacity to modify this intention. And let’s suppose that you also do not have it in
your power to change your relevant means–end belief since the evidence is obvious and
overwhelming. So consider an inference from (R) and:

(1*) You intend E and believe M is a necessary means to E, and these attitudes of
yours are not psychologically modifiable by you
to:

(2) You have a practical reason to intend M.

As Setiya and others have in effect noted (though in somewhat different ways), it is going to
be difficult to insist that in this case we cannot detach the reason-conclusion, (2) (Greenspan
1975; Setiya 2007; Schroeder 2009). After all, in this case, in contrast with the original case,
the only route to conforming to the reason in (R) that is possible for you, given your psycho-
logical incapacity (and the overwhelming evidence for the relevant beliefs), involves
intending M. Intending M is in this fairly strong sense necessary for conforming to the
reason cited in (R). But if we must concede that (2) follows from (R) and (1*) then we
have a problem, since such an inference to (2) is in tension with our claim that intentions
do not in general provide reasons.

One might hold out the hope that even here a Broome-like rejection of detachment will
save the day. After all, in this case, intending M is only psychologically necessary – not
metaphysically or logically necessary – for conforming to the reason cited in (R). But
it seems to me that this is likely to be over-demanding as a general, substantive view
about what is needed for the transmission of reasons along lines of necessity; so I think
we do best to develop our theory in a way that allows for some such detachment of a reason.

I can now state my third preliminary point. Our rejection of the myth theory will need to
involve a defense of something close to (R) as a way of defending:

(a) the distinctive, non-instrumental normative practical significance in the particular
case of means–end coherence.

But we need to do this without:

(b) sanctioning the inappropriate bootstrapping that is threatened by an inference from
(R) and (1*) to (2).

Setiya believes that cases of psychologically non-modifiable intentions show that providing
for (a) ineluctably leads us to (b). This leads him to cognitivism – in contrast with (a) – as
the way to avoid a myth theory. Here again we have the idea – one that also seemed to be at
work in Wallace’s discussion – that the only viable alternative to a myth theory is a form of
cognitivism. And that is an idea that I mean to challenge. In particular, I want to see if we
can have a version of (a) without (b), and thereby be in a position to reject the myth theory
without a retreat to cognitivism.

3.

So let me describe a strategy that promises to yield a version of (a) without (b). The strat-

egy is to cite a reason that makes something close to (R) true, but a reason that is blocked –
at least normally – by psychological non-modifiability of the intention in favor of the end.
This would mean that in those cases in which the intention in favor of the end is psychol-
logically non-modifiable there is not – at least not normally – the associated reason in favor
of means–end coherence, and so there is no danger of detaching a reason for intending the means. In contrast, in normal cases of a modifiable intention in favor of the end there is the cited reason in favor of intention coherence, but Broome’s rejection of a simple ‘factual detachment of a reason’ remains in force.

For such a strategy to work, we need a link to a distinctive, non-instrumental practical reason for means–end coherence of intention, but a reason that normally has as a background condition the psychological modifiability of the intention in favor of the end. What could this reason be?

To solve this problem in our normative philosophy, we need to do some more work in the philosophy of action. Let’s reflect, in particular, on the metaphysics of self-governance.

What is it for me to govern my own practical thought and action? To some extent following Harry Frankfurt, I take it that such self-governance consists at least in part in the guidance of practical thought and action by practical attitudes that speak for the agent and constitute, as Frankfurt puts it, ‘where . . . the person himself stands’ (Frankfurt 1988, 166). As we might say: the ‘self’ governs when relevant attitudes that constitute where the person stands appropriately guide relevant thought and action. Talk of ‘self’ in talk of ‘self-governance’ is not a way of referring to a special entity, ‘the self’ – an entity that sits behind the lens of conscious reflection – but rather a way of alluding to forms of unified psychological functioning that involve guidance by attitudes that speak for the agent and so whose guidance in part constitutes governance by that agent.

Now, I said earlier that the general modes of practical thinking characteristic of planning agency help to constitute and support temporally extended self-governance. The idea, in part, is that a fundamental way in which planning agents like us take a stand – a stand of the sort that is involved in temporally extended self-governance – is to go beyond various conflicting desires and concerns and settle on coherent and consistent intentions, plans, policies, and the like. I do not say that this is a metaphysically necessary feature of self-governance. Rather, the structure of self-governance poses a design problem: how can the psychic economy provide for the kind of practical standpoint whose guidance constitutes the agent’s governance? And, for planning agents like us, the solution to this design problem involves, at bottom, planning structures (Bratman 2007). Indeed, that planning structures help provide such a solution is among the reasons we have to (continue to) be planning agents. (Bratman Forthcoming a)

And now the point to note is that if I – a planning agent – intend E, but also intend what I believe to be incompatible with E, or do not intend what I believe to be necessary means to E, then, barring special considerations, there is no clear answer to the question of where I stand with respect to E. In intending E, I seem to include E within my practical standpoint; but in intending something I believe to be incompatible with E, or in failing to intend believed necessary means to E, I seem to exclude E. In this local area of my psychology, there is a structure of planning attitudes whose functioning in the guidance of my thought and action does not unequivocally either favor or reject E. So with respect to E, there is no clear place where I stand.

Suppose, then, that we consider the possibility of self-governance by a planning agent with respect to possible end E. Such self-governance requires that the agent have a relevant practical standpoint. But we have seen that if the agent intends E but fails to intend believed necessary means to E, there is no clear fact of the matter about what his relevant standpoint is. Means–end coherence of relevant intentions of a planning agent is a necessary constitutive element in her having a relevant standpoint, and so a necessary constitutive element of her self-governance with respect to end E. (And a parallel argument will be available concerning intention consistency.)
This is a point about the metaphysics of self-governing planning agency; it is not yet a
claim about reasons. But the next point is that it is normatively plausible that we normally
have a reason to govern our own lives, though this reason will vary in strength in ways that
depend on the significance of the practical issues at stake. I am inclined to see this as
a reason that is internal (Williams 1981) in the sense that it is grounded in our common
(I do not claim, universal) concern for governing our own lives. Others might see this as
a candidate for an external reason. And other views of the nature of this reason for self-gov-
ernance are possible. But this is a debate we do not need to have here. All that is needed for
present purposes is the normatively quite plausible idea that there is some such reason for
governing one’s own life.14

The next step is to ask how reasons for self-governance extend to reasons for necessary
constitutive elements of that self-governance. If X is a necessary constitutive element of
self-governance, does the reason for self-governance induce a reason for X?

Not quite. For suppose that you are so psychologically damaged that self-governance is
not a real possibility. Then the fact that X is a necessary (though not a sufficient) condition
of self-governance does not seem to speak in favor of X, given that self-governance is not a
real possibility.

It does however seem plausible to say that a reason for self-governance induces a reason
for a necessary constitutive element of self-governance when but only when self-govern-
ance is itself a real possibility.15 So let us take this as given and put it together with our
discussion of self-governing planning agency.

Well, I have argued that means–end coherence of relevant intentions is a necessary con-
stitutive element of relevant self-governance. And we are now supposing that the reason for
self-governance induces a reason for a necessary constitutive element of self-governance
when self-governance is possible. This leads to the conclusion that, when such self-
governance is possible, the reason for self-governance induces a reason for means–end
coherence of relevant intentions. This reason is distinctive, since it does not depend on
the specific reasons for the specific, intended ends at issue; and this reason is non-
instrumental since it is a matter not of an instrumental means to self-governance but of a
necessary constituent of self-governance. We can express this conclusion as follows:

\[(R^*) \text{ When self-governance with respect to end E is psychologically possible for a}
\text{planning agent, there is a distinctive, non-instrumental reason of self-governance}
in favor of the following: if you intend E and believe M is necessary means to E,
then you intend M.}\]

Return now to a case in which the agent intends the end but this intention is tied to a com-
pulsion in such a way that the intention is psychologically non-modifiable. We do not want
to say that, in such cases of a non-modifiable intention in favor of the end, we can in general
detach from a reason for means–end coherence and arrive at a reason to intend the means.
We can now see that to block this threat we need not resort to a blanket prohibition on
detachment of a reason; we can instead appeal to our account of the relevant reason for
means–end coherence.

After all, the reason of self-governance for intention coherence depends on relevant
self-governance being psychologically possible. And it is, I think, normally true that for
a planning agent to be self-governing she needs to have it in her power to modify relevant
intentions in the light of relevant reflection. But it is precisely this capacity to modify a rel-
evant intention in the light of reflection that is missing in the kind of case highlighted by
Setiya, in which it seems we are threatened with unacceptable detachment of a reason
for the means. And what we now see is that in those cases the reason of self-governance for intention coherence is blocked, since a pre-condition of that reason – relevant modifiability of intention – is not satisfied.

This means that we can accept \((R^*)\) while still blocking unacceptable bootstrapping. We are not in general threatened by an inference from \((R^*)\) and \((1^*)\) to \((2)\), since the non-modifiability cited in \((1^*)\) – a non-modifiability needed to make the detachment plausible (in contrast with a simple ‘factual detachment of a reason’ that, following Broome, we have rejected) – normally blocks a pre-condition for the reason cited in \((R^*)\). So in such cases, there is not a relevant reason of self-governance in favor of means–end coherence, a reason from which can be detached a reason for intending the means.

4.

We have been seeking a path between cognitivism about intention rationality and a myth theory. And I think that \((R^*)\), together with its analogue for intention consistency, is the kernel of such an intermediate position. According to \((R^*)\), there is a distinctive, non-instrumental practical reason in the particular case for conformity by a planning agent to the norm of means–end coherence, so long as the agent is capable of relevant self-governance. This reason will be present even in a case of an akatic (but non-compulsive) end intention. But in such a case, our Broome-inspired rejection of factual detachment of a reason blocks a deductive inference to a reason in favor of necessary means.

Granted, \((R^*)\) makes the cited reason for coherence depend on the psychological possibility of self-governance. But \((R^*)\) nevertheless cites a distinctive, non-instrumental reason for means–end coherence that is present in the particular case so long as the agent has the capacity for relevant self-governance. And I take it that a myth theorist would not be content with the observation that the cited reason for intention consistency and coherence in the particular case has, as a background condition, a capacity for self-governance: the purported myth is, I take it, supposed to be a myth about, \textit{inter alia}, the normative significance of consistency and coherence for agents who have that capacity.

As anticipated, we can also extend this strategy to the norm of intention consistency by appeal to a principle along the lines of:

\[
(R^{**}) \text{ When self-governance with respect to end } X \text{ is psychologically possible for a planning agent, there is a distinctive, non-instrumental reason of self-governance in favor of the following: not both intending } X \text{ and intending } Y \text{ while believing that } X \text{ and } Y \text{ are not co-possible.}
\]

This is supported by the claim about the metaphysics of standpoints, that to have a practical standpoint concerning \(X\) – a practical standpoint of the sort involved in the self-governance of a planning agent – one cannot include as elements of that standpoint both intending \(X\) and intending what one believes to be incompatible with \(X\).

What should we say about cases in which the rationality norm applies but, because of the non-possibility of relevant self-governance, there is not the cited reason of self-governance for conformity to the norm as it applies in the particular case? Suppose that in Setiya’s case the person with the smoking compulsion realizes that his best strategy is to lock the cigarette cabinet. If his compulsive intention to smoke somehow persists, he violates the consistency constraint if he also intends to lock the cigarette cabinet and believes that is not compatible with smoking. This would be a \textit{pro tanto} irrationality. But given that his intention to smoke is not psychologically modifiable, self-governance is, we are supposing,
not possible with respect to that end of smoking; so there is not a reason of self-governance to avoid this irrationality.

But if there is not this reason for avoiding this irrationality in this particular case, why does it matter that he is in this way *pro tanto* irrational? Well, it does not matter in the specific way we have been discussing, one that involves a distinctive, non-instrumental reason of self-governance in the particular case. The non-modifiability of the intention to smoke blocks self-governance with respect to that end, and so blocks the reason of self-governance for relevant intention consistency.

Nevertheless, inconsistency with the non-modifiable intention to smoke remains a violation of a norm that is central to planning agency; and there are powerful reasons for being a planning agent. So there is, in this sense, a breakdown in the proper functioning of relevant planning structures. There is a breakdown even though this is the best the agent can do, given his incapacity, and even though an alternative of a simple and rigid responsiveness to consistency of intention in this case would not have been good practical reasoning.16 This is what lies behind the judgment that there remains here a *pro tanto* irrationality, even though – in contrast with most normal cases – there is not a distinctive, non-instrumental reason of self-governance in this particular case to avoid this particular irrationality. And our ability to say this depends on the point, noted earlier, that what we are proposing is not a reduction of practical rationality to practical reasons, but a fundamental link, a link that allows us to steer a path between a myth theory and cognitivism.

There is more to say (Bratman 2009c), but for now my tentative conclusion is that we can indeed chart a course between a myth theory and cognitivism, a course that does not retreat to only a transmission-of-reasons principle. To chart this course, we need to reflect on certain structures of agency and on what those structures do. Here, as elsewhere, our normative philosophy is enhanced by an enriched philosophy of action.

**Acknowledgements**

This essay derives from a talk I gave to the Stanford University Philosophy Department in February 2009. I thank the participants for their many helpful queries. Special thanks to both Sarah Paul and Bruno Verbeek for helpful written comments.

**Notes**

1. The norms of intention rationality that are my concern here are central to the planning theory I develop in Bratman (1987). I have recently been revisiting these matters in the light of both a rich current literature and my own work on self-governance (Bratman 2007). This has issued in a quartet of essays: Bratman (2009a; 2009c; Forthcoming a; Forthcoming b). In the present essay, I try to connect some of the dots. In doing this, I draw substantially from ideas developed in these other essays, especially Bratman (Forthcoming a; 2009c).

2. Raz (2005) and Kolodny (2007; 2008a). Raz writes that ‘there is no *distinctive* form of rationality or of normativity that merits the name instrumental rationality or normativity’ (Raz 2005, 24). Kolodny, who takes himself to be extending Raz’s approach, rejects the ‘myth [that there is] a set of principles that enjoin formal coherence as such’ (Kolodny 2008a, 390). Though there are some complexities here, I take it (a) that in both cases the underlying ideas lead to a rejection of the general claim that intention coherence and consistency have, in the particular case, their own distinctive, non-instrumental normative significance, and (b) that the philosophical interest of such a myth theory lies at least in part in this rejection.

3. That is, the fact that our planning psychology tracks such consistency and coherence thereby makes practical sense to us and is thereby resistant to revision. Though I cannot discuss this here, appeal to such thoughts about the distinctive normative significance of intention consistency would also be an element in a defense of my argument against the ‘Simple View’ in...
response to an important challenge from Hugh McCann. See Bratman (1987, chap. 8) and McCann (1991).
4. For this use of the term ‘cognitivism’ (in contrast with its use in metaethics) see Bratman (1991).
5. In Bratman (Forthcoming a) I discuss the significance of this point for the approach taken in Wallace (2001).
6. I put to one side whether we should accept Wallace’s interpretation of Korsgaard’s views.
7. This may be clearest in Wallace (2006, 119–20).
8. A view along these lines is suggested in Hussain (unpublished).
9. A full story would also need to do justice to Bernard Williams’s insights (Williams 1981) about the relation between normative reasons and the explanation of action.
10. A full story would need to explain in a deeper way why there is this complex relation between intentions and reasons. For present purposes, though, I simply take this as a given.
11. Setiya (2007, 661). Talk of ‘compulsive’ is mine, not Setiya’s; but his remark that ‘there is nothing I can do to change my intention to smoke’ seems naturally interpretable in this way. In any case, I am assuming that the presence of the sort of incapacity at issue here would not be ensured by the mere fact that an intention is caused, or even causally determined, by antecedent conditions.
12. As John Broome has emphasized in correspondence.
13. Here I draw from ideas in Bratman (2009c) where this strategy is developed and defended in more detail.
14. The reason I have in mind is a reason for a garden-variety kind of self-governance that can be present in a natural, causal order. Further, if we do see the reason for self-governance as an internal reason, and we see the concern for self-governance as, while common, not necessary for agency, then we will need to allow for cases in which there is not this internal reason. For present purposes, however, I will suppose that the agents we are talking about do care about governing their own lives.
15. What if relevant self-governance is a real possibility but as a matter of fact the agent is not going to achieve it on this occasion? It seems to me that the reason for self-governance still induces a reason for a necessary constitutive element of the (possible) self-governance. I may here be in disagreement with John Broome, who writes: ‘Suppose you ought to buy a can of paint and decorate your kitchen… But suppose you are not going to decorate your kitchen… Then it seems obvious that there may be no reason for you to buy a can of paint.’ Broome, however, goes on to say the this ‘pattern’ – namely, ‘You ought to $F$; logically necessarily, if you $F$, you $G$; so there is a reason for you to $G$’ – ‘is nevertheless defensible’ (Broome 2007, 176) My reaction to Broome’s example is to agree that an inference to ‘you ought to buy a can of paint’ is questionable; but the inference with which I am concerned is an inference about reasons (given that the end is attainable). (In this respect, the principle alluded to in my text may be more plausible than an analogous principle of Setiya’s concerning what he calls the ‘all-in practical “should”’[Setiya 2007, 660].)
16. I discuss the significance of this last point for assessing an important argument of Kieran Setiya’s in Bratman (Forthcoming b).

Notes on contributor
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. His books include Intention, Plans, and Practical Reason (1987), Faces of Intention: Selected Essays on Intention and Agency (1999), and Structures of Agency: Essays (2007).

References
Intention Rationality


Bratman, Michael E. Forthcoming b. Setiya on intention, rationality, and reasons. *Analysis Reviews*.


Hussain, Nadeem. Unpublished. The requirements of rationality (version 2.4).


