The Policy-Based Approach to Identification

Christian Miller
Wake Forest University
millerc@wfu.edu

In a number of recent papers, Michael Bratman has defended a policy-based theory of identification which represents the most sophisticated and compelling development of a broadly hierarchical approach to the problems about identification which Harry Frankfurt drew our attention to over thirty years ago.[1] Elsewhere I have argued that hierarchical accounts like Frankfurt's cannot avoid well-known problems with higher-order regresses, and that as a result we should adopt a broadly value-based model in their place.[2] Since Bratman motivates his own view with objections to value-based approaches, and since his positive hierarchical proposal is both interesting and important in its own right, it is certainly worth considering the viability of his account in some detail.[3]

In what follows, I summarize the bare essentials of Bratman's view, and then raise doubts about both its necessity and sufficiency. Finally I consider his objections to rival value-based models, and find those objections to be less compelling than he makes them out to be. Unfortunately, given limitations of space, the development of a positive alternative to Bratman's account will have to wait for another occasion.

1. Identification and the Two Standard Approaches

Discussions of identification are typically prompted by the intuitively plausible results of certain forceful examples. Thus in Frankfurt's now famous case, an addict unwillingly ends up satisfying his desire for drugs despite having vehemently resisted the control it has over him

(1971). And in David Velleman's interesting example, a person is surprised to find his voice rising and his temper flaring during a long-anticipated meeting with an old friend towards whom he has slowly and unknowingly been accumulating grievances in his mind for years (1992, p. 126). In these and other well-known cases,[4] the agent in question has somehow failed to identify with the mental states responsible for issuing in her activity, and as a result has thereby failed to identify with the activity itself. In fact, it may even turn out that she is alienated from the causal antecedents of her behavior, and thus from her perspective on the world it is no longer *her* acting but rather something else with which she has not aligned herself.

Much of the literature on identification has focused on trying to better understand the various mental relations we can bear to our *desires*. The following three relations have been discussed most often:

(i)   From the first person perspective, an agent:
         (a) Reflects upon a given desire (or its intentional object).
         (b) Endorses that desire (or its intentional object).
     In such a case, the agent *identifies* with that desire.[5]

(ii)   From the first person perspective, an agent:
         (a) Reflects upon a given desire (or its intentional object).
         (b) Rejects that desire (or its intentional object).
     In such a case, the agent is *alienated* from that desire.

(iii) From the first person perspective, an agent:
         (a) Does not reflect upon a given desire (or its intentional object).
     In such a case, the agent is a *wanton* with respect to that desire.

Note that these three options are not exhaustive; after all, an agent may reflect upon a given desire, and yet be confused about whether to endorse or reject it. Nonetheless, these will be the three central relations for our purposes in this paper.

By way of illustrating these relations, we can return to the two examples mentioned at the start of this section. Frankfurt's example serves as a helpful case of desire alienation; the addict

has a strong desire for drugs but nonetheless has reflected on that desire (or perhaps has reflected more directly on the pursuit of drugs itself) and has rejected it. In Frankfurt's own words, the unwilling addict, "hates his addiction and always struggles desperately, although to no avail, against its thrust" (1971, p. 17). Velleman's example, on the other hand, is an interesting case of someone who is a wanton with respect to a certain part of his mental life; various mental states seem to have been at work subconsciously in the person's mental life prior to the long-anticipated meeting with the old friend, and much to his surprise suddenly come to the surface when the meeting finally occurs.[6]

Clearly any substantive theory of identification must provide a positive account of what endorsement and rejection amount to, and here hierarchical and value-based approaches have emerged as the two most promising strategies. Bratman himself provides a helpful catalog of the main features of the former approach when it comes to desire endorsement:

(i) The agent has a second-order attitude towards the desire.

(ii) The second-order attitude is a conative attitude which constitutes a commitment on the part of the agent such that he or she is appropriately settled on the role of the first-order desire.

(iii) The second-order attitude is concerned with certain kinds of forward-looking functioning of the first-order desire, and includes this guidance of the first-order desire in its own functioning (2003a, p. 224).[7]

For example, according to Frankfurt's original 1971 proposal, desire identification involves the presence of what Frankfurt called a second-order volition, or higher-order desire that the first-order desire in question effectively motivate the agent (1971, p. 16).

Valuational or value-based accounts of desire identification, on the other hand, typically require some kind of *cognitive* endorsement of the normative acceptability of the desire's intentional object.[8] Such cognitive endorsement could, for instance, take the form of a belief

about the goodness of that towards which the desire is aimed. By way of illustration, consider someone who is actually a *willing* rather than an unwilling user of drugs. On the valuational approach, we could understand such a person as having a first-order desire to take drugs together with a belief that his taking drugs is good or valuable. Similarly, an *unwilling* addict could simply be someone with the same first-order desire together with a belief in the wrongness of his drug use. Note that second-order desires might still have a role to play in particular value-based proposals, but only if they are appropriately grounded first in the agent's relevant evaluative norms.

2. Bratman's View

Bratman's policy-based approach to identification has undergone a series of refinements to the point at which it is now a rather complicated view. Here is his core thesis:

(BR) An agent S identifies with desire D iff S has a non-instrumental higher-order self-governing policy, with which she is satisfied, in support of D's functioning, by way of that very policy, as end-setting for practical reasoning (2002, p. 77).[9]

Fortunately we can safely ignore many of the nuances of Bratman's account; all we need for our purposes is Bratman's appeal to the notions of policies, end-setting, and satisfaction.

Bratman explicitly follows Frankfurt in adopting a hierarchical theory of identification. Admittedly, whereas Frankfurt initially appealed to second-order volitions, Bratman instead thinks that we need to make use of what he calls self-governing policies (2002, p. 76). But as far as their ontology is concerned, such policies are still second-order pro-attitudes with a world-to-mind direction of fit. More precisely, they are second-order intentions, where to be an intention is to be a mental state which is irreducible to some combination of beliefs and desires.[10] Furthermore, what distinguishes those intentions which serve as policies is that their contents are in an appropriate sense *general*, since they concern "one's deliberation and deliberative action in

repeated occurrences of situations of certain types" (2001, p. 322).[11] Examples of *self-governing* polices might include, "developing and supporting a strong concern with honesty in writing, or of trying to be more willing to be playful or less inclined to be impatient with others, or of trying not to be so attracted to chocolates or to other temptations" (2000a, p. 47-8).[12]

But a self-governing policy can exhibit a variety of different kinds of concern for intentional attitudes or courses of action, and so according to Bratman a policy-based approach to identification needs to specify the precise way in which such policies are supposed to endorse mental states, and in particular first-order desires. One non-starter is to say that the policy supports treating the relevant desire as issuing in action. For as Frankfurt noted long ago,[13] an agent can identify with two desires while recognizing that he can only act on one of them (Bratman 1996, p. 195-196).[14] So instead Bratman quite reasonably seizes on a plausible alternative, namely that the policies at issue support the agent's treating her first-order desire as providing a justifying *reason* in motivationally efficacious practical reasoning.[15]

Naturally enough, such a move raises worries about circularity, since after all one might think that for an agent to treat a desire as a justifying reason requires that she first have identified with it. Hence we get Bratman's talk of 'end-setting':

> A desire for [end] E *functions as end-setting for practical reasoning* when that desire motivates by way of a process of practical reasoning that appeals to E as a justifying end (2002, p. 75, emphasis his).[16]

Thus according to Bratman, a desire can function as end-setting in this way even though the agent does not reflectively endorse that functioning (2001, p. 323, 2002, p. 74). For example, participants in Stanley Milgram's famous shock experiments could be understood as desiring to continue their participation out of norms of obedience, even though they would not endorse such desires and their corresponding actions. For them, their desire to participate functioned as end-

setting, but they did not endorse and hence identify with that desire (2001, p. 321, 2002, p. 74-5).[17] The fact that end-setting and identification can come apart modally is thus supposed to eschew any worries we might have about circularity.

So a desire's functioning as end-setting is not sufficient on Bratman's view for an agent to identify with such functioning. But he thinks that we can get much closer to sufficiency if we build such functioning into the very *content* of a higher-order self-governing policy whose concern then would be to oversee the role of the first-order desire in the agent's practical reasoning.

Finally, Bratman is well aware that a hierarchical theory of identification must show how to avoid regresses of higher-order pro-attitudes.[18] After all, it seems that an agent can be alienated from a second-order attitude just as much as she can from a first-order one, and merely appealing to a still higher-order state only serves to leave open the same possibility of alienation at that level while also increasing the number of mental states at work in the theory. How is Bratman's proposal (BR) supposed to address this concern? Here is where the third important notion of *satisfaction* is meant to play a role. For Bratman thinks that one can identify with a self-governing policy without requiring other second-order or still higher-order attitudes to bear appropriate relations to that policy. Instead, and again following an earlier proposal by Frankfurt (1992, p. 13-14), Bratman claims that it merely needs to be the case that the policy does not conflict with other self-governing policies about which desires to treat as reason-giving that the agent happens to have at the moment.[19]

Unfortunately, despite its impressive sophistication, Bratman's policy-based hierarchical theory seems to be neither necessary nor sufficient for an agent to identify with his first-order desires.[20] Let us take each of these claims in turn.

3. Against the Sufficiency of (BR)

The sufficiency of Bratman's account rests in large part upon the adequacy of his appeal to satisfaction. Recall that satisfaction is characterized rather narrowly as the absence of conflicts with other standing policies. But this seems inadequate when it comes to ruling out a number of other kinds of conflicts which could potentially lead to discord in the agent and her resultant failure to align herself with a particular first-order desire.

One such kind of case involves a conflict between a second-order policy and a particular second-order desire. As we noted above, two essential features of policies according to Bratman are that they must be non-reducible intentions and that their contents must be general. And given what we are told about satisfaction, (BR) ends up being compatible with the additional presence of a strong second-order desire that a particular first-order desire *not* serve as a justifying reason, at least not in this one case. But intuitively a conflict between such a second-order desire and the agent's general policy in favor of treating first-order desires of this kind as end-setting, could lead to the agent's failure to identify with his first-order desire.

Here is an example. Suppose that a married man has a standing first-order desire to remain faithful to his wife, and that this desire is backed up by a general policy which meets Bratman's conditions as outlined in (BR). But one day the man finds himself in an especially tempting situation, and forms both a powerful first-order desire for a particular woman as well as a strong second-order desire (rather than an intention or decision) that his first-order desire to remain faithful to his wife not serve as justifying in his motivationally efficacious practical reasoning, at least not for the moment. Then it seems at least *possible* for him to no longer fully align himself and thereby identify with the first-order desire to remain faithful. In other words, during the sexual affair itself it doesn't seem to be necessarily true that he would experience his

desire for the woman as an alienating force, while at the same time his will is doing what it can to resist his body's actions. And a few hours later, the husband could be realigned with his general policy and his desire to remain faithful. In the meantime, though, it seems possible that the policy could be set to one side as having authority for the agent. Admittedly, the husband may never have been *alienated* throughout from his first-order desire to remain faithful, but identification and alienation are not mutually exhaustive standpoints,[21] and a momentary failure to identify with his first-order desire is all that is needed to demonstrate the insufficiency of (BR).

In the same way, third-order versus second-order conflicts also pose trouble for (BR). An agent might have a second-order policy, but come to desire that she not have that policy towards the relevant first-order desire. Such a third-order desire need not itself be a policy, and it need not be one from which the agent is alienated either. Indeed, the considerations at work in forming the third-order desire could be considerations about which the agent cares a great deal more than those involved in forming the policy. As a result, the third-order desire can make the agent deeply regret a particular policy and wish that it be abandoned or at least significant revised.

Alternatively, rather than desiring that she not have a particular policy, the agent could simply cease to care directly about, for instance, her having a strong concern for honesty in writing, to use one of Bratman's own examples (2000a, p. 47). In general, there needn't be a specific event which transpires in such a way that the agent no longer cares about her policies; she could just wake up one day and find that some of them leave her cold. At other moments, of course, the contributing factors might be quite clearly defined if, for example, the agent is suffering from depression, enervation, exhaustion, listlessness, and the like. Regardless of the causal factors at work, it seems that a failure to care could lead an agent to not fully align herself with such a policy.[22]

Finally, to use a kind of case towards which I am particularly sympathetic, the agent might reflect upon the normative desirability of having such a policy in the first place. Suppose that after much consideration, she forms the belief that having this particular second-order intention is not a good thing, without yet taking the time to consider which opposing policies to adopt. Then if she reflectively acknowledges both the presence of this belief and the continued existence of the intention, it might be that the importance her values have for her leads her to side with her normative judgment and hence causes her to cease to identify with the operations of the standing policy. So long as the considerations at work in forming the normative judgment have greater importance for her than those which were responsible for the policy, any subsequent actions caused by the policy would be instances of weakness of will with which the agent would not align herself.[23]

I do not take the above catalog of counterexamples to the sufficiency of Bratman's account to be exhaustive; still other cases may be available. Taken together, however, what such cases aim to show is that Bratman's narrow construal of satisfaction will have to be broadened considerably in order to account for the possibility of conflicts with both an agent's cognitive and non-cognitive mental states. Yet at the same time, there are limits to how broad such a construal should be; as Bratman himself rightly notes, "[w]e should not require the complete absence of conflict" (1996, p. 200).

Thus the challenge for Bratman is to revise his account of satisfaction in such a way that certain conflicts between self-governing policies and either normative beliefs or higher-order non-cognitive states must not exist in order for an agent to identify with his desires, while also allowing there to be other innocuous forms of psychological conflict with the policy. In addition, such an account of satisfaction will have to separate out the two kinds of conflicts on principled

grounds; after all, using the examples given above it would be rather easy to put together a list of troublesome mental states in an *ad hoc* fashion. While I have no objections to the availability in theory of such an account, I also have a hard time seeing what such principled grounds could be.

4. Temporally Extended Agency

It might be thought that Bratman has already provided the resources with which to address these purported counterexamples to the sufficiency of (BR). In particular, no attention has been paid thus far to Bratman's extensive discussion in several recent papers of temporally extended agency and the Lockean account of personal identity.[24]

In "Two Problems about Human Agency," Bratman nicely describes what I take to be the fundamental aim in discussions of identification, namely that of explaining how certain of an agent's mental states and actions can be said to represent his or her own fundamental outlook on the world and hence have what Bratman calls *agential authority* (2001, p. 312).[25] Any such explanation of agential authority takes as its foil the fact that, "[a]n agent moved by desires of which he is unaware, or on which he is incapable of reflecting, or from whose role in action he is, as we sometimes say, estranged, seems himself less the source of the activity than a locus of forces" (2001, p. 312). The goal, then, is to isolate certain kinds of psychological functioning that allow the resultant actions to have the authority to speak for the agent herself and not just be the product of causal forces in her psychology.[26]

Do policies like those in (BR) have the authority to speak for the agent? Bratman argues that they do once we appreciate the role that intentions in general and self-governing policies in particular can play in our temporally extended agency. As an agent, I see myself as spread out over time so that, for example, I regard the completion of a project as achieved by the very same agent who started the project, namely me. For Bratman, our personal identity over time is best

understood using the familiar Lockean approach of requiring that central psychological connections be maintained over time.

From here, Bratman's argument is straightforward:

(1) The policies in (BR) are at least some of the attitudes which "have it as a primary role to constitute and support Lockean ties of a sort that are characteristic of our temporally extended agency" (2000a, p. 46).[27]

The justification for (1) concerns the functional role of policies and plans, which is at least in large part to coordinate and organize temporally extended behavior through the inducement of the kinds of connections relevant to Lockean personal identity.[28] Next, we need:

(2) The "actor is, and understands herself to be, a temporally persisting agent whose agency is temporally extended" (2000a, p. 46).[29]

Given (1) and (2), we can infer that:

(3) The policies in (BR) at least partially constitute and support the agent herself.

(4) Hence, the policies in (BR) speak on behalf of the agent and so have agential authority.

Less formally, the basic idea is that the policies in question in (BR) allegedly serve to partially constitute and support who the agent is, and so a desire or action which they serve to legitimate also will be something with which *the agent* identifies.

How is this relevant to the purported counterexamples raised in the previous section? If the policies in (BR) really do constitute and support the agent's perspective on the world, then the agent so constituted would not fail to identify with them in the face of opposing desires or normative beliefs which do not play a similar constitutive role. So perhaps (BR) is sufficient after all.

In response, one could raise various doubts about the general Lockean approach to personal identity, and my own view is that these doubts are serious enough to prompt us to look

elsewhere.[30] Nor has Bratman himself given us any arguments for why we should adopt such an approach.

Even in the context of the Lockean framework, though, I am not convinced that the above response would work. For we need to distinguish between two explanatory claims:

(E1) The agent is aligned with the relevant policies *because* they partially constitute and support the Lockean ties which sustain that agent over time.

(E2) The relevant policies partially constitute and support the Lockean ties which sustain that agent over time *because* the agent has aligned himself with them.

Bratman seems committed to (E1), and it is this claim which is central to his general defense of (BR)'s sufficiency.

Before considering (E1), it might be wondered whether (E2) is even a viable alternative. After all, if like Bratman we are working within an event-causal framework and hence cannot appeal to a metaphysically substantive agent who can align himself with various mental states, then on what basis would an agent come to align himself with the policies in (BR)?[31]

Here, though, is where the valuational model can appeal to its own distinctive resources. Thus, for example, it might be that alignment with a particular policy is secured through a realization of the normative desirability of having such a policy. As a result of being normatively validated, such a policy could then come to constitute and support the ties which connect the agent in the ways described by the Lockean approach.[32]

How then do we decide between (E1) and (E2)? My view is that the intuitive force of the examples presented in the previous section militates against (E1). Here it is important to be clear about the dialectic. The previous examples purported to show that on intuitive grounds alone there is conceptual space between the policies in (BR) on the one hand, and the agent himself and his identification with those policies on the other. Suppose that they do indeed show that

there is such conceptual space. Then since the examples were stated in a way that is neutral with respect to competing theories of personal identity, they can be used as independent evidence against any view which maintains that such policies partially *constitute* the Lockean ties which bind the agent over time. In other words, they can be used as independent evidence against (E1).

Bratman does acknowledge that there are cases in which it is coherent for an agent to have a self-governing policy as in (BR), but there still be an open question by the agent's own lights as to where *she* stands (2000a, p. 60-1). Such cases, Bratman claims, can have the following structure:

> Perhaps the agent finds herself newly impressed with considerations that have not yet been articulated by her as policies or quasi-policies. Perhaps the agent suspects that these considerations may suggest revisions in her self-governing policies and/or quasi-policies. The agent is wondering whether she should continue to be committed to these policies and quasi-policies or whether she should, instead, make changes in them. This is a perfectly coherent thought (2000a, p. 60).[33]

While I certainly agree that such a thought is perfectly coherent (as the examples from the previous section tried to suggest), such an admission seem puzzling for Bratman to make. After all, he claims that (BR) is *sufficient* for the relevant self-governing policy to have agential authority and hence constitute the agent's own outlook. But as these cases involving new considerations show, *the agent* can adopt reflective distance from certain policies, which would be unintelligible if she were partially constituted by them. Nor does the satisfaction condition help, since as Bratman explicitly notes the considerations in question are not other policies.

While Bratman thus makes the surprising admission that such cases are coherent, he thinks that what is *not* coherent is the following thought: " 'I do not want to, or see any reason to change this current package of policies and quasi-policies. And I recognize that they clearly reject D. I just want to ask whether I should let them determine where *I* stand on the present occasion with respect to D' " (2000a, p. 61).[34] But this is also surprising. For I might one day

think about one of my self-governing policies, not because of any newly discovered considerations which bear on the policy, but simply out of curiosity about whether it is a *good thing* for me to continue to maintain the policy. It could be that after further reflection I conclude that it is a good thing, and don't give the matter any thought again. But during that time of reflective examination, it surely *seems* to be a coherent question that I am considering, namely whether *I* should continue to align myself with this policy, and hence whether *I* should continue to endorse or reject the first-order desire which is the concern of the policy. Thus there certainly seems to be conceptual space between myself and my view of the world on the one hand, and certain of my policies on the other. Any view which says otherwise seems to need much more argumentative support than Bratman has given in order to show that this question is not in fact coherent.

The upshot of this section, then, is that given the residual conceptual space between the policies in (BR) and the agent himself, Bratman has yet to solve the problem of explaining agential authority.[35]

5. A Final Challenge to Sufficiency

Let us conclude this critical discussion of the sufficiency of (BR) by returning to Bratman's characterization of end-setting:

> A desire for [end] E *functions as end-setting for practical reasoning* when that desire motivates by way of a process of practical reasoning that appeals to E as a justifying end.

Such a characterization was developed as a result of Bratman's claim that an account of desire identification can't just concern the motivational efficacy of a desire; rather it must also concern the normative standing of that desire's intentional object. Naturally as a defender of a valuational

approach to desire identification, I welcome such a result. But in my view, such an approach brings with it the need for separate normative *beliefs* to be operative as well.

The relevant question here is the following: on what basis does the agent form a policy which in part involves treating E as a justifying end?[36] Presumably there must be some other mental state or states which serve as such a basis. For if an agent has a policy concerning a desire for end E, but is also at a loss as to *why* E should be treated as a justifying end, then the agent can experience disorientation and confusion about such a policy and the role it is playing in his practical reasoning. And such confusion is typically regarded in the literature as being incompatible with genuine identification.[37]

Treating E as justifying involves regarding E as in some way normatively desirable. So where in the agent's psychology could this normative standing for E come from? It doesn't seem that it could result solely from prior desires or intentions that the agent has which are relevant to E. After all, the propositional content we read off from our desires and intentions is purely *descriptive* content like *that I become a philosopher* or *that I vote in the election*. As such, then, it is hard to see how such content could have justificatory force for the agent and sway him on *normative* grounds towards certain courses of action and away from others. And the same holds even for more philosophically loaded contents like *that I maximize pleasure* or *that I act in accordance with what my fully rational self wants my less than fully rational self to want*.[38]

Similarly, it does not seem that E's being regarded as normatively desirable could be the result of most of the beliefs that an agent might have. Such descriptive propositional contents as *that it is raining outside* or *that my parents are still alive* do not seem to be sufficient by themselves to lead an agent to treat an end as justifying.

But there is an exception. If through a process of deliberation an agent arrives at a belief like the following:

(*)   My belief that E is good.

then the content of such a belief naturally could explain why the agent formed a policy which crucially involves treating E as normatively justifying in his practical reasoning.[39]

The implication of the above is that without adding a prior normative belief to the story in (BR), Bratman's account again may not be sufficient. And yet adding such a belief would constitute a significant concession to (if not outright acceptance of) the opposing valuational approach to identification.[40]

6. The Necessity of (BR) and Disjunctivism

Let us leave issues about sufficiency behind for the moment and turn to the question of whether (BR) is necessary for desire identification.[41] Bratman himself concedes that there will be pressure to countenance "cases in which, by the agent's own lights, the evaluative judgment does fully settle the issue of ownership of relevant desires, and on the basis of this the agent is fully behind, or fully rejects, those desires" (2003a, p. 239). Thus I might have a passing whim, come to recognize that there is nothing of positive value and much of negative value to pursing the object of that whim, and so to the extent that I can, reject its influence on my subsequent behavior.

How then does Bratman propose to handle what look to be a sizable number of what he calls *value-judgment-determined* cases (2003a, p. 239)? His response takes the form of an argument from illusion. By showing that we will need to appeal to higher-order pro-attitudes and not just value judgments in certain cases of malfunction and underdetermination – in particular, cases involving weakness of will and ties between competing values – Bratman thinks that this

will make it 'plausible to infer' that we also need such attitudes in the remaining cases as well (2003a, p. 239).

Let us leave aside for now Bratman's arguments for why we need pro-attitudes and not just value judgments in particular instances of malfunction and underdetermination; we shall see later that those arguments are less than convincing. But even *granting* that he is right about such cases, Bratman still needs a way to block standard *disjunctivist* responses to arguments from illusion. In general, disjunctivism denies that the phenomenon in question can be constructed out of an element which is had in common in both defective and veridical cases; rather the phenomenon is understood disjunctively as being either just the defective element or that element which makes the remaining cases veridical.[42] Thus as far as Bratman's argument is concerned, the disjunctivist can simply allow the requisite pro-attitudes into one disjunct of an account of identification (for cases involving malfunction and underdetermination), while letting value judgments do all the work in the other.[43]

Bratman seems to be aware of this problem, but his response underestimates the force of the disjunctivist position. For he thinks that it is enough to meet the disjunctivist's challenge if he can reject the claim that, "it is only in value-judgment-determined cases that there is, as it were, full-blown desire ownership or rejection. The other cases I have emphasized are possible, but as cases of desire ownership or rejection they are, so to speak, second (or, in breakdown cases, third) best" (2003a, p. 240). But we don't even have to look at the reasons he proceeds to give for rejecting this claim in order to see that such a rejection is simply orthogonal to the matter at hand. For the disjunctivist can happily grant that *both* kinds of cases are instances of full-blown desire ownership or rejection while at the same time maintaining that only *one* of them need involve higher-order pro-attitudes of any kind. And given the intuitive plausibility of taking the role of

value-judgments to be sufficient in a wide number of cases for an agent to identify with his desires, it is not at all clear what new arguments Bratman could provide in order to directly address the disjunctivist's challenge.[44]

In fairness to Bratman, he could readily admit that he has not adequately ruled out a disjunctive account of desire identification, but still insist that the burden of proof rests squarely with the advocate of this alternative view. After all, any form of disjunctivism comes out worse compared to its non-disjunctive rival on the grounds of explanatory elegance, as distinct kinds of entities are appealed to in each disjunct of the account.[45] How important this desideratum is in assessing competing accounts of desire identification is difficult to evaluate, but nonetheless when other things are equal it may serve to tip the balance in favor of Bratman's own account.

I am willing to acknowledge that there is something to this response, and furthermore that the defender of the necessity of a value-based theory of desire identification should focus primarily on responding to Bratman's treatment of the malfunction and underdetermination cases in his argument from illusion. Having said this, however, two final points still seem worth making. First, while Bratman's non-disjunctive account comes out ahead with respect to explanatory elegance, it may do significantly worse than the disjunctivist with respect to parsimony. Whereas Bratman is committed to the presence of his highly complex policy state in *every* case of desire identification, the disjunctivist can get by without it in all but a few instances of malfunction and underdetermination.

Secondly, given the manifold complexity and diversity of human deliberative processes, it is not at all unreasonable to think that different agents might come to identify with their desires using rather different mental states. Some might structure their lives around a very rigid and organized set of values, whereas others might give a great deal of leeway to their occurrent

desires.[46] What this suggests is that the burden of proof might actually rest with Bratman and not with the disjunctivist when it comes to thinking that *all* reflection about our desires has to involve the presence of a token of the same specific mental type of policy in (BR) in order for desire identification to occur.[47]

My own view is that value-based theories do not need to avail themselves of a disjunctive formulation since Bratman has not given us convincing reasons for thinking that normative beliefs are either not necessary or not sufficient for desire identification. But suppose that I am wrong about the force of Bratman's objections. The point of this section has been to suggest that a value-based theorist should then take seriously the possibility of falling back on a disjunctive formulation of her view, since Bratman has yet to successfully rule out such a view and since there are several considerations mentioned above which seem to recommend it.

7. Purported Counterexample: Weakness of Will

Bratman describes several families of alleged counterexamples to the sufficiency of value-based alternatives to hierarchical theories of identification and alienation. Given the acknowledged plausibility of a value-based account in a significant number of cases, and given the role that these counterexamples are meant to play in grounding an argument from illusion for Bratman's rival view, clearly a great deal of weight rests upon their proving to be successful.

The first family of examples concerns familiar instances of weakness of will. Here Bratman's argument is primarily the following:

> Perhaps I think it strictly better to be a person who forgives and turns the other cheek but nevertheless, in a kind of self-indulgence, allow into my life a willingness to express reactive anger. Though this role of my desire to express my anger diverges from my relevant evaluative judgments, it is not a desire I reject or disown (2003a, p. 227).[48]

While the case is admittedly under-described, on the face of it there doesn't seem to be a problem here for any sophisticated value-based account. In fact, there are at least four ways that

the valuationist can handle such cases; the choice as to which of these four is preferable can be made only once we have a more richly specified example before us.

First Response. One natural way of understanding the talk of 'allowing into my life' is in terms of a revision of the agent's initial evaluative norms concerning forgiveness in such a way that they build in special exception clauses for certain forms of reactive anger. In this way, the agent's values *do* end up licensing the desire in question, at least in this one instance.

Second Response. Alternatively and perhaps more naturally, it might be that the agent's *moral* norms prescribe forgiveness, but in cases of reactive anger the agent's *egoistic* or self-interested norms have higher priority in his mental life than do his moral norms, and thus he identifies with his angry desire in virtue of its content having been deemed normatively acceptable by this other set of evaluative standards instead. This response accords nicely with Bratman's claim that the allowance for reactive anger was made as a 'kind of self-indulgence.'

Third Response. Suppose, though, that the agent has arrived at what by his own lights is an all things considered judgment against reactive anger. Then it could still be true that he cares very little for the norms operative in forming that judgment, and as a result does not align himself with the judgment. This lack of identification with the normative belief could allow the reactive anger to have free reign without being 'disowned' by the agent. In fact, the agent could not only fail to be alienated from the anger; she may even be able to get herself into a position where she thinks that she is fully behind it. After all, even if a judgment that the anger is not desirable was formed, the agent can exclude this belief from bearing on her anger by her own lights through various acts of self-deception.[49]

Fourth Response. Eventually, though, the value-based theorist will have to draw the line somewhere, and the appropriate place seems to me to be when the following conditions obtain:

(i) The agent forms the all things considered judgment that this expression of reactive anger is undesirable.

(ii) The agent identifies with and hence is aligned with the norms operative in forming that judgment.

If both conditions obtain, then the valuationist has good reason to deny that there are cases in which expressions of reactive anger are not rejected or disowned.

8. Purported Counterexample: Competing Values

The second family of purported counterexamples to the sufficiency of a valuational approach to identification admittedly does receive far more detailed discussion from Bratman. Here is the first such case:

> Suppose that you and I both like to drink alcohol. We both think that there is some value in drinking alcohol when that is what one wants to do. We have both reflected also on the impact of such drinking on our lives, and we both see that there would also be a certain value in systematically abstaining from alcohol. So far, we are alike. But suppose that you go on reflectively to reject your desire to drink alcohol whereas I do not. In what does your rejection consist? The answer offered by a Frankfurtian theory is that it consists in a higher-order, conative attitude . . . this Frankfurtian attitude is to some extent grounded in your judgment of the value of abstinence, but it goes beyond that judgment. You arrive at this higher-order attitude, but I do not; yet we share the cited judgments of value. The fact that your lingering desire to drink is one you disown does not consist solely in those judgments, but essentially involves a Frankfurtian attitude (2003a, p. 227-8).[50]

Does such an example show that the valuational approach can't tell a sufficient story about desire alienation?

One point to make is that it is consistent with the valuational approach that an agent be alienated from a desire in virtue of first forming a normative belief about its undesirability, and then having that belief directly cause the formation of a higher-order Frankfurtian attitude such as a self-governing policy. So long as the attitude is, to use Bratman's language, 'grounded' entirely in the normative belief, it is clear that it is the belief and not the higher-order attitude which is responsible for the agent's alienation from that first-order desire.

Bratman's claim, however, is that in the above case the Frankfurtian attitude 'goes beyond' the judgment about the value of abstinence, and hence has a distinctive role of its own to play in securing alienation. Should we accept this? It is hard to know what to think here since at least three important details have been left out of the example:

(i)   The strength of the first-order desire for alcohol in each of our lives.

(ii) The relative importance in each of our lives of the norms which favor abstinence as compared to the norms which permit drinking.[51]

(iii) The degree to which each of us is disposed to conform our desires to our evaluative outlooks, as opposed to letting our desires revise our norms.

Thus, for example, norms which favor abstinence might play a more central role in your life than mine, perhaps because they derive from prior norms about which you care a great degree more. As a result, while you also might see some value in drinking, it is clear to you that abstinence is all things considered highly desirable. Thus, you might reject your desire to drink and identify with your desire to refrain, whereas I might do the opposite because of the prior difference between us captured by (ii).[52]

Alternatively, suppose that our desires and value judgments are roughly similar, such that we both desire drinking more but initially value abstinence higher than drinking. Then you might judge that your particular desire to have a beer is normatively unacceptable, and given that you are disposed to conform your desires to your evaluative outlook, you reject this particular desire. I, on the other hand, may be such that I am prone to self-deception about what I really value (along the lines of the third response above), or am such that I let my strong desire for a beer allow for a momentary exception cause to be built into my norms (along the lines of the first response above). Either way, I end up accepting rather than rejecting my desire for the beer.

It might be thought that I am avoiding the most sympathetic reading of Bratman's case, namely that we should understand the features mentioned in (i) through (iii) as being roughly equal in both of us, and yet it also is true that one of us rejects whereas the other accepts a desire to drink. In this version, then, it surely must be that one person's acceptance and the other's rejection is grounded at least in part in the higher-order Frankfurtian attitude.

Thus let us take up underdetermination cases in our final section.

9. Purported Counterexample: Underdetermination by Values

The third and final class of counterexamples on offer to the sufficiency of a valuational approach to desire identification concerns cases where the agent is (i) faced with genuine alternatives (such as drinking or abstinence) and yet (ii) does not form a comparative normative judgment about which alternative is better. Bratman mentions three variants of such cases:

(a)   The alternatives are equal in value.

(b)   The alternatives are such that the agent is uncertain as to how to rank order them.

(c)   The alternatives are of no comparative value (2003a, p. 230-1).

But, Bratman says, it is also true that many times, "life must go on. One way for him to settle the issue would be to settle on a [higher-order Frankfurtian pro-attitude]" (2003, p. 230).

Bratman's general point here is that there are cases in which, "the agent's value judgments by themselves underdetermine his stance in response to the practical issues raised about how he is to live" (2003a, p. 231). What should the valuationist say in response?

As an aside, it is worth noting that underdetermination cases may cut against hierarchical theories as well. As Bratman himself acknowledges, settling on a higher-order pro-attitude is only *one* of several ways whereby an agent might respond to failed attempts at rank ordering the relevant values (2003a, p. 230-1). Nor does Bratman claim that an agent's availing herself of one

of these alternative means would preclude her from identifying with the subsequent actions. As a result, it is difficult to see why higher-order pro-attitudes would be *necessary* in such cases, and hence difficult to see why the hierarchical approach in general and Bratman's policy-based approach in particular can justifiably purport to be offering necessary conditions for desire identification.

More importantly, there still seems to be an important role for normative judgments to play in underdetermination cases. For when faced with a lack of clear guidance from his values, an agent needn't immediately choose one of the relevant options over the others; he may instead prefer to spend time gathering further information which could serve to tip the normative balance in favor of a particular alternative. Alternatively, he may not care very much about any of the alternatives and simply disregard them altogether. Thus the valuationist might argue that it is only when the agent values *settling* the issue of which course of action to pursue, that he proceeds to avail himself of one of the means at hand for deciding in such a way that he can subsequently identify with the resulting action. Regardless of whether a higher-order Frankfurtian attitude is formed or not as a means in order to help arrive at a particular resolution, it is the value that the agent places in *resolving* the underdetermination which on this proposal serves to ground his alignment with the relevant desires and actions which are subsequently formed.

Thus I conclude that even in the cases Bratman takes to most strongly illustrate the need to supplement a valuational account of identification with higher-order pro-attitudes such as those found in (BR), it is by no means clear that such a move is actually required. And without it, his argument from illusion cannot even get off the ground.

10. Conclusion

In this paper I have tried to raise a number of serious problems for Michael Bratman's recent work on identification. Bratman's view represents the most sophisticated version of a hierarchical theory of desire identification, and if the objections above are indeed compelling, then perhaps it is time for us to rethink the merits of the hierarchical approach in general. My own view is that we should instead examine whether a value-based theory can provide an illuminating alternative to the hierarchical models which have dominated the discussion of identification for so many years.[53]

Works Cited

Bratman, Michael. (1987). *Intention, Plans, and Practical Reason*. Cambridge: Harvard University Press.

_____. (1996). Identification, Decision, and Treating as a Reason. *Philosophical Topics*, 24. Reprinted in Bratman 1999, pp. 185-206.

_____. (1999). *Faces of Intention: Selected Essays on Intention and* Agency. Cambridge: Cambridge University Press.

_____. (2000a). Reflection, Planning, and Temporally Extended Agency. *The Philosophical Review*, 109, 35-61.

_____. (2000b). Valuing and the Will. *Philosophical Perspectives*, 14, 249-265.

_____. (2001). Two Problems about Human Agency. *Proceedings of the Aristotelian Society*, 101, 309-326.

_____. (2002). Hierarchy, Circularity, and Double Reduction. In Buss and Overton 2002, pp. 65-85.

_____. (2003a). A Desire of One's Own. *The Journal of Philosophy*, 100, 221-242.

_____. (2003b). Autonomy and Hierarchy. *Social Philosophy & Policy*, 20, 156-176.

_____. (2004). Three Forms of Agential Commitment: Reply to Cullity and Gerrans. *Proceedings of the Aristotelian Society*, 104, 327-335.

Buss, Sarah and Lee Overton (Ed.). (2002). *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge: MIT Press.

Cullity and Gerrans. (2004). Agency and Policy. *Proceedings of the Aristotelian Society*, 104, 315-325.

Dancy, Jonathan. (1995). Arguments from Illusion. *The Philosophical Quarterly*, 45: 421-438.

_____. (2000). *Practical Reality*. Oxford: Oxford University Press.

Frankfurt, Harry. (1971). Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*, 68. Reprinted in Watson 2003, pp. 322-336.

_____. (1977). Identification and Externality. In A. Rorty (ed.), *The Identities of Persons* . Berkeley: University of California Press. Reprinted in Frankfurt 1988, pp. 58-68.

_____. (1987). Identification and Wholeheartedness. In F. Schoeman (Ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*. Cambridge: Cambridge University Press. Reprinted in Frankfurt 1988, pp. 159-176.

_____. (1988). *The Importance of What We Care About*. Cambridge: Cambridge University Press.

_____. (1992). The Faintest Passion. *Proceedings and Addresses of the American Philosophical Association*, 66, 5-16.

_____. (1999). *Volition, Necessity, and Love*. Cambridge: Cambridge University Press.

_____. (2002). Reply to Michael E. Bratman. In Buss and Overton 2002, pp. 86-90.

_____. (2004). *The Reasons of Love*. Princeton: Princeton University Press.

McDowell, John. (1982). Criteria, Defeasibility, and Knowledge. *Proceedings of the British Academy*, 68, 455-479.

Milgram, Stanley. (1983). *Obedience to Authority*. Perennial.

Miller, Christian. Impossibilities of the Will. Manuscript.

_____. Meaning and Caring. Manuscript.

_____. The Structure of Instrumental Reasoning. *Philosophy and Phenomenological Research*. Forthcoming.

_____. (2004). *Agency and Moral Realism*. Ph. D. Dissertation. University of Notre Dame.

Olson, Eric. (1997). *The Human Animal: Identity without Psychology*. Oxford: Oxford University Press.

Snowdon, Paul. (1980). Perception, Vision, and Causation. *Proceedings of the Aristotelian Society*, 81, 175-192.

Stump, Eleonore. (1988). Sanctification, Hardening of the Heart, and Frankfurt's Concept of Free Will. *The Journal of Philosophy*, 85, 395-420.

Thau, Mike. (2003). What is Disjunctivism? *Philosophical Studies*, 193-253.

Velleman, J. David. (1989). *Practical Reflection*. Princeton: Princeton University Press.

_____. (1992). What Happens When Someone Acts? *Mind*, 101. Reprinted in Velleman 2000a, pp. 123-144.

_____. (2000a). *The Possibility of Practical Reason*. Oxford: Clarendon Press.

_____. (2000b). Introduction. In Velleman 2000a: 1-31.

Watson, Gary. (1975). Free Agency. *The Journal of Philosophy*, 72. Reprinted in Watson 2003, pp. 337-351.

_____. (1987). Free Action and Free Will. *Mind*, 96, 145-172.

_____ (Ed.). (2003). *Free Will*. Second Edition. Oxford: Oxford University Press.

[1] Bratman 1996, 2000a, 2000b, 2001, 2002, 2003a, 2003b, and 2004. For Frankfurt's influential work on identification, see the papers in his 1988 and 1999.

[2] See my *Agency and Moral Realism*, Ph.D. dissertation, University of Notre Dame.

[3] Surprisingly, despite the number of papers that Bratman has devoted to his view in recent years and despite what I take to be the novelty and importance of his account, there has been very little discussion of it in the literature. Cullity and Gerrans 2004 represents, so far as I am aware, the only extant critical discussion, and briefly suggests two counterexamples to Bratman's view, neither of which is related to claims made in what follows here. Bratman responds to Cullity and Gerrans in his 2004.

[4] See in particular Frankfurt 1977, p. 63, 67 and Bratman 1996, p. 196.

[5] Compare Bratman 1996, p. 186, 2000a, pp. 38-39, 2000b, p. 256, and 2002, p. 66. There is some debate in the recent identification literature about whether the second condition should be an endorsement condition or merely a weaker acceptance condition. For relevant discussion, see Frankfurt 2002, p. 87. Bratman himself repeatedly uses the language of endorsement in the passages cited above, but nothing in what follows should hang on the outcome of this debate. Similarly, there has been much discussion of whether a theory of desire identification should involve the agent endorsing her desires themselves or merely the content of those desires. See in particular Watson 1975 and Bratman 2003b. Again, we can remain neutral on this debate for our purposes here.

[6] As Velleman writes, "We can assume that this causal relation was mediated by any number of subconscious intentions – intentions to sever the friendship by alienating my friend, to alienate my friend by raising my voice, to raise my voice now . . . etc." (1992, p. 126 fn. 12).

[7] To save space, I have condensed Bratman's catalog of five theses down to three.

[8] For such a view, see in particular Watson 1975. Also helpful here is Bratman's characterization in his 2003a, p. 223. Alternatively, a valuational account might require that the belief concern the normative acceptability of the *desire* rather than just the desire's intentional object. While hierarchical in structure, such a view would not be classified as a hierarchical approach to identification given Bratman's schema since the higher-order mental state is a cognitive *belief* rather than a conative attitude like a desire or intention.

[9] See also Bratman 2000a, p. 55.

[10] For plausible reasons in favor of the irreducibility of intentions to combinations of beliefs and desires, see Bratman 1987.

[11] See also his 2000a, p. 41, 48, 2000b, p. 263 fn. 27, 2002, p. 83 fn. 34, 2003a, pp. 225-226, and 2003b, p. 160.

[12] See also Bratman 2000a, p. 41 and 2003b, p. 160.

[13] See Frankfurt 1977, pp. 66-67 and 1987, p. 170.

[14] A similar problem arises for a related specification of the content of the agent's policy in terms of its support for the desire's functioning as an effective motive. For related discussion, see Bratman 2000a, pp. 52-54, 2000b, pp. 257-358, and 2002, pp. 67-68.

[15] This is a brief summary of Bratman 1996, pp. 196-198, 2000a, pp. 52-55, 2000b, p. 258, 2001, p. 321, 2002, p. 67, and 2003a, p. 225.

[16] See also his 1996, p. 198 and 2001, p. 323.

[17] For the experiments themselves, see Milgram 1983.

[18] This was the main problem for Frankfurt's earlier view. For related discussion, see Watson 1975, Frankfurt 1987, p. 166, and Bratman 1996, p. 188 and 2000a, p. 37.

[19] See Bratman 1996, pp. 200-201, 2000a, p. 49, and 2002, p. 77. Bratman refines this characterization in his 2000a, p. 50, but the objections which pertain to Bratman's satisfaction condition in what follows also apply equally well to his revised characterization.

[20] Bratman explicitly construes his proposal as offering necessary and sufficient conditions in his 1996, p. 202, 2000a, p. 39 fn. 12, and 2002, p. 77. Compare also Cullity and Gerrans 2004, p. 317. Emphasis on the purported necessity of the view can be found in his 2002, p. 78. Emphasis on the purported sufficiency of the view can be found in his 2004, p. 327.

[21] As Bratman himself agrees (1996, p. 197).

[22] For much more on the nature and role of caring, see Frankfurt 2004 and my "Meaning and Caring."

     In his first paper on identification, Bratman did seem to appreciate the force of this kind of case, and added the requirement to his account of desire identification that the agent not only decide (i.e., form a policy) to treat the desire as reason-giving, but also *actually* treat it as reason-giving, or at least be prepared to do so in the relevant

circumstances (1996, p. 202). Curiously, though, this requirement seems to have dropped out of Bratman's theory in the course of refining this initial account in subsequent papers. And perhaps for good reason, since Bratman would then owe us much more about what is involved in bridging the gap between *actually treating* a desire as reason-giving as opposed to merely deciding or forming a policy to do so.

[23] We shall discuss weakness of will in more detail in section seven. For an attempt to capture the sense of importance at issue in the above text, see my "Impossibilities of the Will."

[24] See especially Bratman 2000a, as well as his 2000b, pp. 256-257, 2001, p. 319, 2002, p. 76, and 2004, p. 328.

[25] See also his 2000a, p. 39, 2003a, p. 222, and 2003b, pp. 168-169. More precisely, the aim is to provide such an explanation while also remaining in an event causal framework and hence eschewing reference to agents themselves as substances with causal powers. For relevant discussion, see Velleman 1992 and Bratman 2000a, pp. 39-40 and 2001, p. 312.

[26] See Bratman 2001, p. 312, 320 and 2003b, pp. 168-169.

[27] Bratman explicitly affirms (1) on pages 47-48. See also his 2000b, p. 257, 2001, p. 322, and 2002, p. 76.

[28] See Bratman 2000a, pp. 47-48, 2000b, p. 257, 2001, p. 322, 2002, p. 76, and 2003b, p. 169.

[29] See also Bratman 2002, p. 76.

[30] For some important doubts and an attractive alternative, see Olson 1997.

[31] For more on this framework, see note 25.

[32] It is worth emphasizing that this is only one suggested value-based approach, and needn't be the only one or the best one available.

[33] For quasi-policies, see note 41.

[34] See also his 2004, p. 333.

[35] Of course, a rival valuational approach needs to be able to tell its own story about why and in what conditions a given set of norms can have agential authority, and how that authority can then be transmitted to the desires which those norms endorse. The story is long and complex, and so in the interest of space I take up the task of providing such an account elsewhere. See in particular my "Impossibilities of the Will."

[36] Bratman himself briefly mentions this question in his 2001, p. 323, but quickly sets it aside.

[37] See in particular Velleman 1989, 1992, and 2000b.

[38] I take up the issues here in much greater detail in my "The Structure of Instrumental Reasoning."

[39] Again the issues here are complex, and the interested reader is encouraged to read the paper in the previous note for a more detailed treatment.

[40] For a view which features higher-order pro-attitudes grounded in prior evaluative beliefs, see Stump 1988. In his 2003a, Bratman seems to agree that, "Frankfurtian attitudes will normally be to some extent grounded in and constrained by reflection on what one takes to be of value" (226). So one might think that here Bratman explicitly concedes that (BR) is not sufficient for desire identification. However, as we will see below, Bratman goes on to consider cases of malfunction and underdetermination in which he claims that Frankfurtian higher-order attitudes alone settle the issue about desire identification, and then uses those cases in an argument from illusion to show that even in 'normal' cases in which the Frankfurtian attitude is grounded in a normative judgment, it is still the attitude which is responsible for ownership.

[41] It is worth noting that at times Bratman claims that (BR) is not necessary given that (i) there are other higher-order intentions which are similar in certain important ways to the self-governing policies in (BR), but (ii) they are also relevantly different so as to constitute mental states distinct from self-governing policies, and yet (iii) these intentions can secure first-order desire identification. Bratman calls such intentions self-governing 'quasi-policies' (2000a, pp. 57-60).

However, we can set this complication aside since my concern is whether intentions *of any kind* are necessary and sufficient for desire identification, and so what follows also applies to an account which appeals to both self-governing policies and quasi-policies.

[42] Here I have been helped by Dancy 1995 and 2000, pp. 138-145. The classic papers defending disjunctivism in the theory of perception are Snowdon 1980 and McDowell 1982. See also the extensive discussion in Thau 2003.

[43] Remember that Bratman's arguments do not attempt to *directly* show that all value judgments will be insufficient for identification. Rather, they try to show the insufficient of value judgments in cases involving malfunction and underdetermination, and then *generalize* from this result to the remaining cases of action. The disjunctivist alternative, on the other hand, accepts the insufficiency of value judgments in the first two kinds of cases, but denies the generalization and hence denies that value judgments are insufficient in the remaining cases.

[44] This is not to say that it isn't important for Bratman to *also* be able to reject the proposal quoted above. Rather the point is that such a rejection is largely irrelevant when it comes to responding to disjunctivism.

[45] Thus in the case of desire identification, Bratman's account has explanatory elegance in the sense that it can appeal to the role of the same kind of higher-order pro-attitude in all cases of genuine identification with a first-order desire, whereas the disjunctivist would have to posit the role of such pro-attitudes in cases of malfunction and underdetermination while allowing cognitive value judgments alone to do the work in at least many of the remaining cases. This latter position, then, in relying on such different kinds of mental states to secure desire identification, has the appearance of being explanatorily disjointed and hence inelegant. I have benefited from discussion of this issue with Don Hubin.

[46] I try to suggest ways in which a valuational approach can allow for such psychological variation in my "The Structure of Instrumental Reasoning."

[47] More precisely, the burden of proof might rest with Bratman when it comes to thinking that all reflection about our desires has to involve the presence of a token of a self-governing policy in (BR), *or* a token of a self-governing quasi-policy. For the latter, see note 41.

[48] Bratman's example derives from Watson 1987, p. 150. See also Bratman 1996, pp. 189-190.

[49] This response points to an important demand which any valuational theory should try to satisfy, namely that of not only explaining how values and norms can secure desire identification, but also of explaining how the agent first comes to identify himself with these particular values and norms. For similar remarks, see Velleman 1992, p. 134 and Bratman 2003a, p. 226. I have attempted to provide such an explanation of what might be called 'norm identification' in my "Impossibilities of the Will."

[50] For the other cases, see Ibid., pp. 228-229 and 2003b, p. 159. In general, these cases are supposed to take the following form:

> In each of these examples there are relevant judgments of value on both sides of a practical issue; and a [higher-order pro-attitude] is to some extent grounded in some of those judgments. But the presence of these evaluative grounds does not entail that there is no further work to be done by the [higher-order pro-attitude]. In some such cases it is only when one arrives at a [higher-order pro-attitude] that one has, in the relevant sense, taken a stand with respect to the issues raised for one's life (2003a, p. 229).

Note that Bratman makes it clear that here he is attacking the sufficiency rather than necessity of value judgments.

[51] For what this sense of importance amounts to, see my "Impossibilities of the Will." The account there does not make use of any higher-order non-cognitive mental states.

[52] Bratman seems to accept something like this response. See his 2003a, pp. 229-230.

[53] For helpful comments, I am grateful to Michael DePaul, Philip Quinn, and several anonymous referees. An earlier version of the paper was presented at the 2004 American Philosophical Association Pacific Division Meeting. I am very grateful to Don Hubin for excellent comments and to Michael Bratman for helpful discussion.