

---

---

## Symposia on Gender, Race and Philosophy

Volume 7, number 1. Winter 2011

<http://web.mit.edu/sgrp>

---

---

### Reply to Benson, Christman, Rocha, and Stoljar

**ANDREA C. WESTLUND**

Department of Philosophy  
University of Wisconsin-Milwaukee  
P.O. Box 413  
Milwaukee, WI 53201  
[westlund@uwm.edu](mailto:westlund@uwm.edu)

I'd like to begin by thanking Sally Haslanger and the editorial board of SGRP for organizing this symposium, and Professors Benson, Christman, Rocha and Stoljar for their very insightful and stimulating comments. I have learned a lot from this discussion and, in particular, from the juxtaposition of my paper with Catriona MacKenzie and Jacqui Poltera's fine article on narrative integration. Some illuminating connections have been drawn between the two, and my reply begins from the ground we share. I proceed thematically, since certain ideas arose in more than one set of comments.

i.

Natalie Stoljar points out that both MacKenzie and Poltera's paper and my own draw a distinction between the necessary conditions for cross-temporal agency and the necessary

conditions for autonomy. Stoljar argues, further, that both papers put pressure on what she calls the "authenticity" criterion of autonomy. Paul Benson makes the same claim in his closing paragraph and, while John Christman seems to retain a stronger connection between authenticity and autonomy than either Benson or Stoljar, he also argues that my view is not best understood as offering an authenticity condition. This emergent theme has prompted me to rethink the way in which I have sometimes framed my argument.

Authenticity conditions, as both Stoljar and Christman note, are usually paired with competence conditions as two distinct components of an account of autonomy. Competence conditions concern capacities one needs in order to be self-governing (rationality, self-control, and the like), while authenticity conditions pertain to the question of what makes a belief or desire truly "one's own." On certain hierarchical views of autonomy, for example, reflective endorsement satisfies the authenticity condition, while the competence condition would be satisfied at least in part by possession of whatever capacities one needs to engage in motivationally effective reflective endorsement. I have been critical of such accounts. Nonetheless, I have sometimes framed my own position as an alternative account of authenticity, or, in other words, as an alternative answer to the question of what's required for a belief or desire to count as one's own. (I do this implicitly in the paper under discussion, and more explicitly in Westlund 2003.) Stoljar suggests that, because of the distinction I draw between autonomy and cross-temporal agency, perhaps I should not frame it in this way at all.

I am amenable to this suggestion. As useful as the language of ownership can be, it also has its limitations. I have never been strongly tempted by the idea that self-governance depends on determination by something we might call one's "true" self. I have found the notion of ownership pertinent to autonomy primarily because it is suggestive of a normative

relationship of responsibility or answerability for the motives on which one acts. Whether those motives involve desires or values that form a part of one's self-identity (or, indeed, one's narratively integrated identity) is a conceptually distinct issue. Such belonging is not, at the very least, sufficient for autonomy. The discussion in this symposium has, I think, convinced me that the notion of authenticity is more naturally and more clearly associated with the question of integrated identity than it is with autonomy. If this is so, then one upshot of my argument (like MacKenzie and Poltera's) would be that autonomy and authenticity can come apart.

ii.

But if responsibility for self is not an authenticity condition, then what is it? Christman suggests that it is best understood as a competence condition. As Christman sees it, what a person lacks, when she lacks the disposition to answer for her commitments, is "the ability to sift properly through the reasons available to her for having any values at all" (Christman 4). Going this route, according to Christman, would preserve the anti-perfectionism of my view – since self-subordination is not directly ruled out – but only at the cost of rendering its relationality derivative. That is, autonomy would turn out to be (constitutively) relational only to the extent that an agent's ability to sift through reasons was itself dependent on her engagement in certain kinds of social relationship. This, Christman argues, may be true of many of us, but not all.

I would argue, however, that answerability must involve something more than the ability properly to sift through reasons. A deeply deferential individual may be quite *capable*, in a perfectly ordinary sense of the word, of examining objections, constructing arguments, and, indeed, answering for herself. What she does not do is hold herself to any expectation that she do so. The self-responsible agent, by contrast, holds herself to the *expectation* or *demand* that she

respond appropriately to legitimate requests for justification, experiencing herself as *owing* a suitable response, at least under certain conditions. One who relates to herself in this way feels the demand for answers not just as a pressure from without, but also as a requirement to which she is held from within. If answerability for oneself is a competence, then, it is a normative one, resembling in certain respects what Stephen Darwall calls "second-personal" competence.<sup>1</sup> The non-answerable agent fails to stand in a distinctive normative relationship to herself, which is just the flip side of her insensitivity to the legitimate justificatory demands placed on her by others.

The normativity of the competence puts a more than merely derivative relationality back in the picture. Stoljar highlights a key passage in my paper, which I'll cite again here: autonomy requires "positioning oneself as always a potential member of a reflective or deliberative dyad," such that "the internal psychological condition of the autonomous agent ... point[s] beyond itself, to the position the agent occupies as one reflective, responsible self among many" (Westlund 35). Answerability, in sum, is normative in the sense that the autonomous agent holds herself to an *expectation* or *demand*, and relational in the sense that an expectation or demand is something that one party is subjected to by another, or by one party to herself when she manifests sensitivity to what others can legitimately ask of her.

iii.

But does this send me from the frying pan into the fire? In embracing a normative conception of competence, have I lost what Christman calls my anti-perfectionist credentials? Christman, I think, would argue that the normative competence I've just described implies a commitment to an ideal of social relations about which reasonable persons could disagree. While Benson does not seem to think I'm committed to that kind of ideal, he does argue that a view of

this sort cannot be purely formal. To hold oneself answerable is, he notes, “to apply to oneself a standard that calls for one to answer to potential criticisms (under certain conditions)” (Benson 3). This, Benson argues, is not a purely formal standard, since such expectations are matters of substantive disagreement.

These are important objections, with which I am still grappling. I would argue, however, that my view falls short of building in a commitment to ideal social relations – or, at least, of doing so in a way that falls afoul of liberal anti-perfectionism. Indeed, my argument is meant to proceed in what Christman identifies as a broadly Kantian way, falling within the tradition on which “moral truth cannot be specified independently of practical reason,” and according to which “we cannot specify the truth about morality and the good life before fully grasping how a reasonable, self-governing agent is structured and functions in the first place” (3). I think (and hope) I have avoided relying on any *prior* commitment to a particular vision of the good life. If our form of practical reasoning and self-governance turns out to have a normative and relational structure, this is, at the least, not something imported from outside.

I would argue, moreover, that there is room for a significant amount of disagreement (even radical disagreement) between autonomous individuals about what forms of social engagement they should embrace. To be self-governing, we don’t need to have a clear understanding of what *makes* us self-governing, nor are we required to endorse the idea that every individual *should* in fact function in whatever way self-governance turns out to require. We just need to function that way in fact. To put it somewhat paradoxically, we don’t need to embrace our own answerability, we just need actually to be self-answerable. A shared moral or political commitment to relationships in which all are treated as answerable will likely provide a more hospitable

environment for the development and sustenance of the relevant competence – but this does not mean that one cannot have the competence without the commitment. (I will return to this point two paragraphs hence and, from another angle, in my final section.)

My response to Benson’s objection, regarding justificatory practice, is similar. In short, I suggest that my account’s commitments on this score are not such as to violate the spirit of a formal approach to autonomy. While my account does seem to imply a commitment to *some* standard governing legitimate challenge, the shape of that norm is as much up for grabs (and as much the subject of justificatory discourse) as anything else. Some people might bristle at the thought that they *ever* owe a response – and yet, whether they hold themselves answerable for that commitment is crucial to the question of their autonomy. I submit that, in this respect, the requirement I defend is no more substantive than the conditions given by other defenders of “formal” views. Reflective endorsement theorists, for example, do not require that the autonomous agent actually endorse the criterion of reflective endorsement itself; reflective endorsement is held to be constitutive of autonomy, regardless of whether the agent would, at some level, fail to reflectively endorse that very requirement.<sup>2</sup>

I realize that the highly deferential but purportedly still answerable agent is bound to strike many readers as a paradoxical creature. Yet it seems to me that we can function as answerable even while being highly uncertain, and indeed confused, about our own status as such, and about the implications of our own reflectively endorsed commitments. Setting aside those cases of deference I describe as “deep,” even someone who is highly committed to deferring to particular others might hold herself answerable for that commitment. I do not mean that she might be able, ultimately, to successfully defend her commitment against all

criticism. But responsibility for self is not in that sense a *success* condition for autonomy. Rather, it is a relational stance on which one holds oneself answerable for one's commitments. Indeed, I take this stance to involve a certain humility about the ultimate defensibility of our commitments, by which I mean openness to the possibility that we may turn out to be *unable* to defend some of them in the end. Without that kind of humility, it is hard to see how one would be genuinely engaged by and responsive to external critical perspectives.

iv.

This reference to humility might seem surprising, particularly given the emphasis placed on self-respect and self-worth by many defenders of relational autonomy, and (within this symposium) by Benson and James Rocha. Humility is surely compatible with self-respect or self-worth, but it suggests a different emphasis. So, I now turn to the important question of the relationship between autonomy and these self-regarding attitudes.

Benson argues that, even apart from the justificatory standards discussed above, other normative conditions "lie submerged in attitudes that are preconditions for agents' holding themselves answerable to others' criticism" (3). In particular, he suggests, the self-answerable agent must regard herself as sufficiently competent and worthy of answering for herself. While these attitudes do not, on his view, commit one to perfectionism – there is plenty of room, he suggests, for non-liberal understandings of self-worth – he does think they render a view like mine at least "weakly" substantive.

I'm inclined to agree that autonomy could be undermined by a severe lack of confidence in one's ability to answer for oneself. (Incidentally, however, I think it could also be undermined by an *excess* of confidence, since this might leave one short on the sort of epistemic humility that I take to be

crucial to genuine answerability.) But a lack of self-regard would, in this way, pose a threat to autonomy on just about *any* formal account, since these seem, quite generally, to rely on competences that could be disabled by severe enough self-doubt. This line of thought leads me to doubt whether there is a strong distinction to be drawn between formal and weakly substantive views of autonomy. The formal views might *all* involve more of substance than we have thought, and there might be nothing at all worrisome about that. What we do need, however, is a distinction between both "formal" and "weakly substantive" views, on one side, and, on the other, the more strongly substantive views that seem to have perfectionist implications. I am convinced, like Benson, that this distinction can be drawn, though I need to think more about whether or to what extent we ultimately differ on the details.

It's possible that we differ in our understandings of the role played by self-worth. Self-worth, of course, is a peculiar and multifaceted thing. As Thomas Hill Jr. and others have noticed, some deferential agents might be quite proud to be self-subordinating. (The Deferential Wife he describes in "Servility and Self-Respect" seems to be one of them.) In one intuitive sense, the proud self-subordinator lacks a sense of self-worth, since she takes her interests to be of secondary importance at best. But in another sense, she does not. She regards herself as having a proper role to play and takes pride in playing it well. There is thus an equally intuitive sense in which her self-worth *depends* on her deference.

I argue that whether deference compromises autonomy depends on how deep it goes. If it goes so deep that the agent does not experience herself as even potentially engageable in justificatory dialogue (because she hears all challenges to her action-guiding commitments as addressed through her to someone else), then her autonomy is undermined. She does not have the normative, relational competence I describe in

Section III above. But, if it does not go that deep, and she is engageable in justificatory dialogue, then her apparently self-denigrating views do not render her non-autonomous. Does this mean that she must have an independently identifiable attitude of self-worth or self-respect that serves as a precursor to the self-relation manifested in her answerability? As my treatment of “paradoxical” agents already suggests, I’m reluctant to take this route: human psychology strikes me as too complex to tie self-answerability to any specific, substantively defined attitude or belief about oneself. I’m more inclined to say that her self-answerability is (as I say in the paper) a *sui generis* form of self-respect: a form of self-respect that is constituted *by* the normative self-relation I describe.

Rocha raises an important worry about the scope of this account, arguing that its lack of substance precludes proper characterization of the harm oppression does to agents in undermining their self-respect. He writes: “The severity of the harm done to these agents lies precisely in the fact that their deeply internalized values *prevent* the full achievement of their autonomy” (Rocha 4). There is something strange, he says, about considering an agent to be autonomous when her values have been warped by “some obscure – but external – process” (Rocha 4). The strangeness to which Rocha refers is what initially got me interested in this topic, and I do agree with Rocha that oppression can incur distinctive, autonomy-related harms. But there is a way in which I can partially accommodate his point. On my view, oppression can undermine both autonomy and (a distinctive form of) self-respect precisely because it can undermine answerability. I describe deference, for example, as deep when it does so. Perhaps, in the paper under discussion, it looks as though I regard this kind of damage as relatively rare, but that is not necessarily the case. It might be more common than we would like to think. Combined with oppression-related damage to self-confidence, which could have the effects noted

above, my net might be cast more widely than Rocha suggests.

There will, however, be cases on which we disagree. On my view, the obscure but external processes at which Rocha gestures are “domesticated” when the agent holds herself answerable for the commitments they help to shape. Even if the agent is harmed by these processes in some other way, she is, at least, not gripped by the commitments they generate: these commitments, and the processes from which they flow, are subject to her review and assessment. Stoljar rightly points out that, in some contexts, occasions to review and assess will be limited. I believe that cultures and traditions are generally complex and interconnected enough to offer at least some footholds for critical challenge. Nonetheless, I am willing to bite the bullet and argue that, even where there is no such foothold, there is an important difference between the agent who is answerable but unchallenged, and the one who is not answerable at all. Both lead lives that are deeply affected by their internalization of oppressive values. But the former has an attitude (which I’ve described as a form of humility) toward those values that renders her open to rethinking them.

I would submit that for most of us, at least some of our deeply held values fall in the “unchallenged” category, and we should not consider those values to be heteronomously held simply because they have never been questioned. How we would respond under counter-factual conditions matters to our autonomy. The Bigsby character described by Rocha, in fact, strikes me as someone who is gripped by his values, rather than self-governing, because he is so extremely (indeed, comically) unable to entertain the possibility that they may be mistaken. He is not self-answerable, ironically, because he is too confident in the arguments that “support” what turn out to be self-denigrating values.

v.

I have not done justice to the four excellent commentaries I have been fortunate to receive on this paper: each raised a wealth of important questions, some of which I have addressed only partially and provisionally, and others, I fear, not at all. All of the comments will continue to inform and challenge my thinking on these matters, and for that I am grateful. So, in closing, I would once again like to thank all for their contributions to this discussion.

---

<sup>1</sup> See Darwall (2006). I have more to say about the relationship between self-answerability and second-personal competence in Westlund, forthcoming.

<sup>2</sup> I thank Nate Jezzi for making this point some years ago in his comments on an early draft.