

THE AUTONOMOUS LIFE: A PURE SOCIAL VIEW

Michael Garnett

In this paper I propose and develop a social account of global autonomy. On this view, a person is autonomous simply to the extent to which it is difficult for others to subject her to their wills. I argue that many properties commonly thought necessary for autonomy are in fact properties that tend to increase an agent's immunity to such interpersonal subjection, and that the proposed account is therefore capable of providing theoretical unity to many of the otherwise heterogeneous requirements of autonomy familiar from recent discussions. Specifically, I discuss three such requirements: (i) possession of legally protected status, (ii) a sense of one's own self-worth, and (iii) a capacity for critical reflection. I argue that the proposed account is not only theoretically satisfying but also yields a rich and attractive conception of autonomy.

Keywords: freedom, autonomy, republicanism, non-domination, self-worth, self-creation, manipulation

1. Introduction

In contemporary debates about the nature of autonomy, one particularly disputed issue is whether or not, and if so to what extent, it ought to be conceived in essentially social or relational terms. On this issue views may be usefully divided into three categories. Members of the first, which we may call *pure nonsocial views*, seek to elucidate autonomy by means of conditions that make no essential reference to other agents [Meyers 1989; Benson 1991; Frankfurt 1999; Friedman 2003; Korsgaard 2009; Christman 2009].¹ Most classic accounts of autonomy, such as those that analyse the notion exclusively in terms of a capacity to reflect in certain ways upon one's own beliefs and desires, are purely nonsocial in this sense. By contrast, some more recent theorists have argued that a convincing account of autonomy must, *in addition* to such nonsocial conditions, also involve some that are essentially social: such as that one not be subject to coercion or deception, or to the dominating power of others [Dworkin 1976; Raz 1986; Oshana 2006; Taylor 2009]. These are the *hybrid views*. Finally, some have suggested that autonomy might be best elucidated in terms of social conditions alone; as simply requiring, for instance, that one be disposed to attempt to justify one's attitudes in dialogue with critical interlocutors [Westlund 2003]. These are the *pure social views*.

¹ Note that the issue here is the extent to which autonomy is *constitutively* social; a view may be purely nonsocial in this sense despite including a thoroughly social account of autonomy *attainment*, as do a number of these accounts.

This paper proposes a pure social view of autonomy, albeit one that differs significantly from existing views of this kind. On the account defended here, an autonomous agent is one that is hard to push around, that has a mind of her own, and that is difficult to use or to bend to one's will. More precisely, it is argued that an agent is autonomous to the extent to which he is resistant to subjection to foreign wills. This view is purely social, since it analyses autonomy just in terms of this one social condition. Despite this, however, I argue that it is able to accommodate many of the nonsocial conditions advanced by pure nonsocial and hybrid theorists. This is because the property of resistance to interpersonal subjection, although itself essentially social, is one that is *conferred* by various base properties that may themselves be either social or nonsocial. Thus I also argue that many of the most common and familiar conditions of autonomy (such as a capacity for critical reflection, a sense of self-worth, noncoercion, and so on) are properties of agents that work to increase their immunity to subjection to foreign wills, and so are autonomy-conferring properties according to the proposed account.²

Moreover, the proposed view differs significantly from existing social approaches to autonomy. Classically, autonomy, or self-rule, has been understood as *rule by the self*, with the self in turn most commonly identified with one's rational or reflective nature. For their part, social autonomy theorists have tended to remain within this theoretical framework, arguing that either reflective rationality or the self or both are in some way irreducibly social.³ On the proposed view, by contrast, self-rule is understood not as rule by the self but as the absence of rule by others. It therefore has no special need for a theory of the self and can be mostly neutral with respect to competing accounts of selfhood [Garnett 2013]. Similarly, it does not treat autonomy as having any necessary connection with critical reflection (although, as we shall see, such reflection nevertheless turns out to be important).

I argue that this view has five main strengths. First, since (as I shall attempt to show) many properties of agents widely thought to be important for autonomy are in fact ways of manifesting the key property of resistance to interpersonal subjection, it is able to accommodate many of the intuitions that motivate competing views. Thus it has the capacity to be what J. S. Taylor [2009: 1–2] calls a 'capturing' account of autonomy. Second, whereas some extant accounts of autonomy present these familiar conditions simply as so many items on a heterogeneous list, the proposed account provides them with explanatory unity by showing how each contributes to the realization of a single property. Third, the view is able to preserve the standard etymologically-informed analysis of autonomy as self-rule without becoming

² The following analogy may be helpful. A *sedative* is (at least in part) something that tends to cause drowsiness. This is a *psychological* property. But it is one that is typically realized by some base chemical property, such as that of being an opiate or a barbiturate or an alcohol, which is itself *nonpsychological*. In the same way, I claim that the concept of autonomy is best analysed in terms of the *social* property of immunity to interpersonal subjection. But this is a disposition that is in turn realized by one or more base properties, such as that of having a capacity for critical rationality or a sense of one's own self-worth, that may themselves be nonsocial.

³ For instance, Westlund [2003: 515] holds that a failure of dialogical critical reflection constitutes a failure of autonomy by way of constituting a loss of the 'distinct and determinate self'.

embroiled in problematic debates about the nature of the self. Fourth and fifth, the view allows for autonomy to be instantiated in ways that are both consistent with other important human goods and of relevance to beings like us.

These claims will, I hope, be vindicated in what follows.

2. The Account

2.1. Subjection to a Foreign Will

Autonomy, I claim, is resistance to subjection to foreign wills. What is it, then, to be subject to a foreign will, in the relevant sense? To work up to the notion at issue, let me first introduce the more basic idea of *being got to do something*. We may say that *A* gets *B* to *x* just in case *B*'s doing *x* is brought about by *A* acting so as to bring about *B*'s doing *x*. Moreover, we may say that the extent to which *A* gets *B* to *x* is a matter of the significance of the causal role played by *A*'s action in the etiology of *B*'s doing *x*. So if *A*'s action makes only a small difference to the likelihood of *B*'s doing *x*, we may be reluctant to assert that *A* has got *B* to *x*; by contrast, if *A*'s action massively raises the likelihood of *B*'s doing *x*, then we are much more inclined to assert that *A* has got *B* to *x*.

With this idea so understood, we can turn next to that of subjecting a person to one's will. In short, subjecting a person to one's will is a particular way of getting that person to do something. Specifically, subjection to a foreign will involves two further conditions, since it requires both that one be *subject* to the will of another and that that will be relevantly *foreign*. Both conditions are relatively familiar, and I consider them briefly in turn.

The first is motivated by the need to distinguish between cases in which *B* freely and independently chooses to act *in accordance* with *A*'s will, for instance out of love or respect for *A*, and cases in which *B* is truly subjected to *A*'s will. Other things equal, you do not subject me to your will simply by asking me to do something for you. Nor, indeed, do you do so by making a trivial threat (or offer), so long as I could still be reasonably expected to stand up to your attempted influence if I so wished. By contrast, were you to make a substantial threat, or a conditional offer of something I cannot do without, then you would likely succeed in subjecting me to your will. Call this *the reasonable resistance condition*: in getting *B* to *x* by means of issuing an offer or a threat, *A* does not subject *B* to her will if *B* believes that, had *B* refused to do *x*, the consequences that *B* would have thereby suffered are not such as to render this refusal unreasonable.

Note that this condition may be interpreted in various ways, depending on how we understand the relevant notion of 'reasonableness'. It is perhaps helpful to think of different versions of the notion as varying along two dimensions. The first concerns the extent to which a version takes account of *B*'s specific, real-world psychological characteristics, as opposed to abstracting away from such particularities. Many 'reasonable person' tests, such as those familiar from legal contexts, lie towards the latter, more impersonal end of this spectrum. Yet since we are here concerned specifically with

the idea of subjection, as opposed to that of moral responsibility, the relevant notion of ‘reasonableness’ in this context is perhaps more likely to be a subjective one (i.e., some version of what Richards [1987] calls the ‘pressure view’; see also Benn and Weinstein [1971] and Benditt [1977]).

The second of the two dimensions is *sensitivity*: just how bad must an alternative be in order to be deemed unreasonable? While relatively insensitive interpretations may set a high bar, deeming an alternative unreasonable only when it involves some extreme hardship, relatively sensitive interpretations may set a low bar, deeming an alternative unreasonable when it involves even a relatively minor setback. For instance, a *very* highly sensitive interpretation might deem it unreasonable to expect a timid worker to stand up to his boss, just on the grounds that he finds the prospect unpleasant. Note that even on such an extreme interpretation the crucial distinction between being *subject* to a foreign will and *freely choosing to act in accordance with* a foreign will remains: the worker who reluctantly does what his boss demands out of fear of confrontation remains distinct from the worker who cheerfully acts out of a genuine desire to help (assuming, of course, that desires for x are not always just identical with aversions to not- x). Clearly, determining exactly how sensitive the relevant notion of reasonableness ought to be for present purposes would require a careful marshalling of intuitions across a broad range of cases with the aim of reaching reflective equilibrium. Since the main argument of the paper does not turn on the specific outcome of this large task, however, I do not undertake it here. Readers may therefore interpret the reasonableness condition in whichever of these ways they deem most plausible.

I turn now to the second of the two conditions on subjection to a foreign will. This one is motivated by the thought that an agent is not subject to a foreign will when subject to forms of influence that he endorses. Thus the will of the hypnotist I hire to help me to quit smoking is not relevantly foreign to me—though it would be if she also illicitly attempted to add a desire to buy her book. Similarly, I am generally happy for people to try to convince me of what they take to be true by presenting me with good reasons for believing it: this is why you do not usually rule me by rationally persuading me of something. If, however, it were clear that I did not want convincing of something—for instance, of the fact that my missing child is probably dead—and you convinced me of it anyway, then you would have subjected me to your relevantly foreign will. Call this *the conformity of wills condition*: in getting B to x , A does not subject B to his will if it is the case that, were B to know that A intends to get him to x by certain means, B would endorse A ’s intention (where this endorsement is not itself a product of B ’s prior subjection to a foreign will).⁴

To subject a person to one’s will, then, is to get her to do something, where neither the reasonable resistance condition nor the conformity of wills condition is met. It may be helpful to distinguish broadly between two ways

⁴ This may seem to render the account viciously circular, but it merely makes it recursive. The impossibility of A ’s having already performed an infinite number of actions means that the chain of analysis will not run for ever: eventually we will reach either an action that was not intentionally influenced by another, or intentional influence that was not endorsed.

in which one may get someone to do something. Consider an agent in a standard choice situation, in which he selects from some set of opportunities on the basis of a set of preferences. One way of getting such an agent to choose a particular option is to meddle with his opportunity set, by removing, adding, or altering his possibilities of action; another is to leave his opportunity set untouched but, instead, to meddle with his preference set, so as to dispose him to select a different option from an unchanged opportunity set. Each of the above conditions may be thought of as corresponding with one of these methods. If one meddles with another's opportunity set, one subjects her to one's will only if one makes it unreasonable for her to opt for anything other than one's intended option. And if one meddles with another's preference set, one subjects her to one's will only if she would not have freely endorsed one's interference were she to have known of it.

This is, then, what it is for *A* to subject *B* to his will. But what is it for *B* to be subject to *A*'s will? This is a subtly distinct phenomenon.⁵ Suppose that your husband is prone to fits of violent rage when his desires go unsatisfied, and that your only reasonable choice is to ensure that they are met. However, he does not aim to get you to do this; he hardly thinks about you at all. So he does not subject you to his will. Nevertheless, you are subject to his will. And it is this latter fact that matters for the proposed account of autonomy. To take the final step to an account of subjection to a foreign will, therefore, we need to weaken the foregoing account such that, in cases where the reasonable resistance condition has application, the requirement is not that *A* intends that *B* do anything but, more simply, that *B* has no reasonable choice but to act in accordance with *A*'s desires.

The final (unavoidably convoluted) account is therefore something as follows. *B* is subject to *A*'s will just in case (i) *A*'s desire that *p* brings it about that *B* attempts to bring it about that *p*; (ii) where the process described in (i) proceeds via some change to *B*'s opportunity set, neither the reasonable resistance condition nor the conformity of wills condition is met; (iii) where the process described in (i) proceeds *not* via some change to *B*'s opportunity set, it is the case both that (a) *A* intends to bring it about that *B* brings it about that *p*, and (b) the conformity of wills condition is not met.

2.2. Resistance to Subjection to Foreign Wills

On the proposed view, an autonomous agent is one that is resistant to subjection to foreign wills. Thus autonomy is not simply a matter of *happening* not to be subject to a foreign will: the easily manipulable person who gets lucky and is never in fact manipulated is nevertheless nonautonomous, on this account.⁶ Rather, an autonomous agent is one whose freedom from external control is counterfactually robust.

Autonomy in this sense is conferred on agents by their possession of various more specific properties. Let me call these *autonomy traits*. An

⁵ Thanks to an anonymous referee for helping me to see this (by means of the following example).

⁶ Similarly, the highly unmanipulable person who gets unlucky and finds herself subject to a foreign will on some specific isolated occasion may nevertheless be highly autonomous.

autonomy trait is any property that confers on an agent an increased tendency not to be subject to foreign wills. Below I suggest that possession of legally protected status, a sense of one's own self-worth, and a capacity to reflect critically upon one's attitudes are all important autonomy traits. Yet given this very broad characterization, it is clear that the list of possible autonomy traits is potentially endless. It is therefore important to see that, in order to count as autonomous, an agent need not possess every autonomy trait. Autonomy should be thought of as a 'threshold' concept: what matters is that we have some sufficient degree of resistance to interpersonal subjection, not that we have the maximum. There is of course room here for substantive disagreement about how much is enough, and different theorists may set the bar in different places. For current purposes the key point is only that autonomy does not require perfect or absolute unsubjectability. And this means, importantly, that no one autonomy trait is either necessary or sufficient for autonomy. Rather, an autonomous agent is simply one who possesses some sufficient number of sufficiently powerful autonomy traits.

The proposed view therefore allows for a variety of instantiations of autonomy, each constructed from some different combination of traits. There may be any number of such instantiations; nevertheless, some may be more appealing or attractive than others. Specifically, there are two dimensions in which autonomy traits may be assessed. One is *relevance*: we are interested in properties that are potentially possessable by beings like us, and that protect our independence in environments like those that we in fact inhabit. For instance, it is (as far as we know!) irrelevant in this sense that possession of an indeterministically free will might protect us from subjection to certain sorts of super-powered manipulators. The other is *value*: some autonomy traits are properties that we have independent reason to value, while others are properties that we actively disvalue. For example, intelligence is an autonomy trait that is of independent value, while some vices, such as certain forms of arrogance, may also be autonomy traits. An attractive instantiation of autonomy, or package of autonomy traits, is one made up of traits that are both relevant to beings like us and independently valuable (or, at least, not disvaluable).

The account of autonomy proposed here is an account of the autonomy of agents themselves, and not of their actions, choices, or mental attitudes. It is thus an account of *global* and not of *local* autonomy [Dworkin 1988: 15–16]. This is significant, since many of the most important applications of the theory of autonomy involve the concept of local autonomy. We want our theory of autonomy to tell us, for instance, whether a patient is likely to make an autonomous choice concerning her medical treatment, whether a contract was autonomously entered into, and to what extent our democratic decision procedures aggregate autonomous preferences. The account of autonomy proposed here is not designed to tell us any of these things. Instead, it is intended to elaborate a notion of autonomous agency understood as part of a more general ethical, political and social ideal [Feinberg 1989: 44–7]. The purpose of this global notion of autonomy is to constitute a possible component of a general account of human flourishing. On many views, part of what it is for a human life to go well is for it to be autonomous. In explaining

what this means—in explaining what it is for a life, as opposed to some particular action, to be autonomous—we must look to a global account of autonomy.⁷

A complete account of the human good would detail every aspect of the good life and indicate how these competing values are to be weighed and traded against one another. There is no reason to think that autonomy should be the supreme or even the most important of these values; a rich human life is most probably not one that sacrifices all other goods at the altar of autonomy. Indeed, far from being the whole of human flourishing, the life of autonomy is not even sufficient for the life of freedom, inasmuch as full freedom requires not only that one be self-ruled but also that one enjoy a rich set of options.⁸ Nevertheless, it is perhaps safe to maintain that autonomy is at least one important part of what it is for a life to go well. For this reason, any conception of autonomy that presents us with a stark choice between it and other values is likely to be an unappealing one. As we have seen, however, the proposed account allows for autonomy to be realized in different ways, and this flexibility permits the articulation of forms of autonomy that fit within our broader theories of the good.

2.3. Pathologies of Autonomy

It may help to illuminate some of the foregoing discussion by briefly considering a pair of unattractive autonomy traits. It is a consequence of the account of subjection to foreign wills presented in §2.1 that we are all, to some degree, subject to the wills of our parents and educators. This should not be surprising: as social creatures, we are never entirely independent of the wills of others. Nevertheless, it is also a consequence of this account that were a person somehow, impossibly, to break decisively with his past by forging a new identity through some spontaneous self-creative act, then his autonomy would indeed be increased. And this, in turn, helps to explain the fascination that some have had for the idea of *self-creation*. This idea, which at the very least necessitates some fairly heavy-duty metaphysics and is even then probably incoherent, has been rightly criticized by many theorists [Dworkin 1976: 24; Taylor 1982; Feinberg 1989: 33–5; Noggle 2005; Oshana 2006: 155; Taylor 2009: 97–102]. Indeed it is clearly delusional to think that anyone could create themselves in this way, let alone that this could be a universal condition of human beings—a delusion sustainable only in the context of a

⁷ Note also that, whereas part of the job of an account of local autonomy is to explain the normative authority of choices typically made by competent adults, there is no similar reason why levels of global autonomy cannot be ascribed to animals and infants. Indeed, it is not unusual for people to say such things as that some intransigent toddler is attempting to assert her autonomy, or that cats tend to be more autonomous than dogs. These sorts of extended global autonomy claims are easily accommodated by the proposed account (and are wholly mysterious on most existing accounts).

⁸ On the distinction between autonomy and the enjoyment of options, see Dworkin 1988: 14, Feinberg 1989: 62–8, and Taylor 2009: 108–9. Note that options may be restricted by ‘internal’ as well as by ‘external’ obstacles: in either case an agent might remain autonomous whilst nevertheless lacking freedom. Moreover, while Joseph Raz [1986: 372] famously includes ‘an adequate range of options’ as a condition of autonomy, this disagreement is largely terminological: whereas Raz uses ‘autonomy’ broadly to include the enjoyment of options, the proposed account uses ‘autonomy’ narrowly to refer just to something like his ‘independence’ condition (while preferring ‘freedom’ as the broader term). For some argument in favour of preferring this latter terminological framework, see Garnett [2013].

governing ideological framework in which the importance and even existence of constitutive social relations in general, and of relations of care specifically, are rendered invisible. Moreover, it may also be the case that there is value in being embedded in an unchosen culture, such that even were radical self-creation possible it would still, all things considered, represent an unappealing ideal.

The pure social account of autonomy need not stand in the way of these familiar criticisms. So powerful are they, however, that it can sometimes seem mysterious how such an idea could have ever, even erroneously, held so much appeal. Yet on the proposed view there is no mystery. Self-creation is an autonomy trait: it is indeed true that, were we capable of self-creation, we would be more autonomous than we are. Nevertheless, self-creation is an *unappealing* autonomy trait. It is irrelevant, since it is not a property that beings like us are likely to possess. And it has disvalue, since its exercise would require the sacrifice of other important goods, such as social belonging. There are simply other, more appealing ways of achieving autonomy.

Similar considerations apply to a second trait, that of radical self-sufficiency. For some, the ideal of autonomy is manifested to the highest degree in the life of the ‘rugged individualist’, who turns his back on society to live a life of isolated self-reliance. This ideal has also been rightly criticized [Jaggar 1988: 27–50; Nedelsky 1989; Held 1993: 43–63, 174–91], though again it is important to diagnose its errors correctly. In particular, the problem is not that a desire for self-sufficiency has nothing at all to do with autonomy: to the contrary, it is quite comprehensible why a person with an overriding concern for autonomy might see some attraction in such a life. The problem, instead, is that such a form of autonomy is so very costly in terms of other important values (such as those of love, community, solidarity and friendship). In turning his back on society, the rugged individualist ignores the ways in which we may live autonomously within society, and the ways in which the right kinds of social relations can increase rather than decrease autonomy.

Self-creation and radical self-sufficiency are both examples of unattractive traits that do nevertheless bear genuine connections to the idea of autonomy. We might therefore think of the tendencies to ascribe significant value to properties like these as *pathologies of autonomy*: evaluational malfunctions to which many of us are prone. Part of the strength of the account under consideration is its ability to explain the relevance of such properties to autonomy whilst also endorsing their critiques. The task now is to describe how autonomy may be realized by traits that are both relevant and valuable. In the next section I discuss three: having the right kind of social and legal status, having a robust sense of self-worth, and having a capacity for critical reflection.

3. Towards An Appealing Form of Autonomy

3.1. *Social and Legal Status*

One of the best ways to make oneself difficult to manipulate is to occupy the right type of social position in the right type of social order—or, at least, to

avoid occupying the wrong type of social position in the wrong type of social order. Consider, for instance, an occupant of an extreme case of the latter: the slave. As the legal property of another, the slave has little say in how he behaves, acting mostly on the basis of commands backed up by standing threats of physical violence. Or consider a wife in mid-Victorian England, also lacking in important legal protections. Were she unfortunate, such a woman's situation could be one approaching that of the slave: forced simply to obey the unreasonable commands of her husband on pain of physical violence, with little realistic possibility of redress. These are central cases of subjection to foreign wills. They are also, intuitively, cases of impaired global autonomy.

On the proposed social account of autonomy, it is one's liability to be subject to foreign wills, and not one's actual subjection to them, that diminishes one's autonomy. So even a slave with a benevolent master—say, one who has no interest in owning slaves, but has inherited one and yet feels under pressure not to free him, and so leaves him largely to his own devices—is, on this account, importantly deficient in autonomy. This is because, although his master does not now subject him to his will, he nevertheless retains the power to do so. Whatever leeway the slave enjoys is dependent upon his master's inclinations. And given that the law supports his master's right to do with him what he will, he is every bit as *susceptible* to his master's control as a more typical and less fortunate slave.

This thought—that freedom or autonomy is a matter of counterfactually robust independence from the control of others—is essentially a republican one. Yet while the social account proposed here incorporates something like the republican conception of freedom, it need not be committed to the republican's more substantive political claims. On Philip Pettit's [1997] view, for instance, the proper role of the state is to maximize non-domination by minimizing the extent to which its citizens participate in dominating relationships. Against this, anti-perfectionist liberals typically argue that the state should be neutral with respect to the good and allow individuals to enter into whichever private relationships they choose. As regards this dispute, however, the proposed account of autonomy takes no position. While it describes a substantive normative ideal, it brings with it no view concerning the proper role of the state with respect to the ideal. In particular, the account has nothing to say about whether the state should license or forbid choices in favour of reduced global autonomy.⁹

Moreover, not only is the present account of autonomy uncommitted to republican ideas about the role of the state, but it also extends and develops the basic republican way of thinking about freedom in significant ways. In the republican tradition the idea of unsubjectability is typically elaborated

⁹ Indeed, the account allows that choices in favour of reduced global autonomy may themselves be autonomous. Thus a Western woman may exercise perfect (local) autonomy in choosing to marry a Saudi and relocate to his homeland, despite the fact that this is, given the social and legal status of women in Saudi Arabia, a choice in favour of reduced (global) autonomy. In so far as we are bound to respect one another's autonomous choices, we are bound to respect this woman's choice of life, even though the life she autonomously chooses is one of relative nonautonomy. (For this reason the present account is not vulnerable to the criticisms levelled by Christman [2004], which are better construed as objections to social accounts of *local* autonomy. See also Westlund [2009].)

in relatively narrow terms, focused on immunity to illegitimate forms of coercion, as opposed to illegitimate forms of control more generally (including, for instance, more subtle forms of psychological manipulation), and also on institutional means of achieving such immunity, as opposed to more personal means. One of the motivations behind the view proposed here is that, by expanding the basic republican idea to include these broader categories, we might find a theory that gives explanatory unity to the whole range of otherwise apparently unrelated conditions that are commonly deemed necessary for autonomy.

3.2. *Self-Worth*

A second property of importance for autonomy is possession of a sense of one's own self-worth. Paul Benson [1994] illustrates the relevance of such a trait with reference to the 1944 film *Gaslight*. In this story a man, Gregory, attempts to prevent his new wife, Paula, from discovering his criminal activities by leading her to believe that she is losing her mind. He makes her think that she is performing nonsensical actions and then forgetting that she has done so, that she is losing things, and that she is delusional. The result is that Paula—initially a fairly strong person—is reduced to a state of confusion and disorientation whereby she no longer regards herself as a competent agent.

For Benson as well as other theorists [Govier 1993; Oshana 2006: 81–3], this undermining of Paula's trust in her own competence constitutes a reduction in her autonomy. Moreover, this is an intuition endorsed by the present account, since a sense of self-worth is an important autonomy trait: as Gregory well knows, those with a low sense of self-worth are easier to push around and to manipulate. However, understanding exactly how and why self-worth functions as an autonomy trait requires a careful look at the property.

According to Benson [1994: 660], 'the sense of worthiness to act which is necessary for free agency involves regarding oneself as being competent to answer for one's conduct in light of normative demands that, from one's point of view, others might appropriately apply to one's actions.' Yet there is something potentially misleading about this formulation, since a person with a very low sense of self-worth may have a warped view of what normative demands may be appropriately applied to his actions. Benson writes [ibid.: 662]: '... the gaslighted woman, who believes that she is losing some portion of her sanity, will be likely to distrust her competence to answer for her conduct in relation to any normative domain.' Yet, for precisely the same reason, she will surely deny that it is appropriate for others to demand that she answer for her conduct.

For present purposes, I instead understand a person's sense of worthiness to act simply in terms of that person's confidence in her own competence as an agent. A person with a healthy sense of self-worth is one who believes that she is, *in general*, just as capable of forming true beliefs and making good decisions as anyone else. So when someone with a reasonable sense of

self-worth encounters someone with a belief contrary to one of her own she may accord *some* weight to this; she may take it, for instance, as a reason to review her own evidence. But if her reasons for believing as she does still appear convincing, the *mere* fact that someone else has reached an opposing conclusion should not, other things equal, cause her to revise her belief.

By contrast, a person with a weak sense of self-worth is one who, in the same situation, will tend to abandon his belief. This is because such a person deems himself less epistemically competent than others, and so (reasonably, given this) judges the prior probability of his belief being correct to be much lower than that of the other's belief being correct. That is, even if he initially takes himself to have good reasons for believing that *p*, the mere fact that another believes that not-*p* will cause him to lose confidence in his assessment of these reasons and hence in his belief. Moreover, these points about belief apply in equal measure to judgments about what is valuable or about what some person has most reason to do.¹⁰

Those with low self-worth are therefore more likely to allow others to alter their practical judgments irrespective of their own assessments of the reasons. Of course, if such a person is lucky enough to be surrounded by others who genuinely care for her and always seek to influence her in ways that she judges to be for her own good, then, on the present account, she will not *in fact* be subject to a foreign will. Yet she will nevertheless be deficient in autonomy. For if she is unlucky, then it will be relatively easy, as in the case of Gregory and Paula, for unscrupulous others to manipulate her in ways that serve their interests at the expense of her own. Thus a lack of self-worth entails an increased susceptibility to subjection to foreign wills, and hence a reduction in autonomy.

It is indeed no accident that many manipulators, and abusive partners in particular, seek to damage their victims' sense of self-worth as a way of rendering those victims more susceptible to their effective control. Nor is it an accident that systems of political domination are often backed up by systems of ideological domination that work to lower the self-worth of the dominated party. Going back to the case of slavery, note that no system of slavery would be practically feasible were it based on threats of physical violence alone: it would be far too costly for any slave master to constantly monitor and discipline any more than a tiny handful of slaves in this manner. Instead such institutions are typically sustained by various auxiliary measures, including ideological systems that work to break slaves' spirits by convincing them that they are unfit to be anything other than slaves [Benson 1994: 658–9].

Finally, it is worth noting that while some may suffer from a deficiency of self-worth, others may suffer from an excess of it. A person with a wildly inflated sense of self-worth, who judges herself to be far *more* competent than

¹⁰ Self-worth is therefore here understood as a type of self-confidence or self-trust. It is more specific than *self-esteem*, which involves a more comprehensive self-assessment [Sachs 1981]: one may judge oneself to be a bad person in various respects without doubting one's competence as an agent. It is also distinct from *self-respect*, understood as knowledge of and appropriate concern for one's own moral standing [Hill Jr. 1973]. There is no conceptual reason why a person cannot fail to claim his own moral entitlements while nevertheless regarding himself as a perfectly competent agent. This said, however, there are no doubt many reasons why self-worth, self-esteem and self-respect should tend to wax and wane together.

anyone else, will be unlikely ever to revise her judgments in the light of the advice of others, no matter how compelling the reasons given. Such arrogance represents both a rational and an ethical failing. Nevertheless, in so far as it tends to make a person even less susceptible to the control of others, it functions as an autonomy trait. Clearly, however, it is an unappealing autonomy trait, and unlikely to form part of an instantiation of autonomy that fits within a broader view of human flourishing. A reasonable and healthy sense of one's own self-worth is, by contrast, both relevant and independently valuable, and an important part of an attractive form of autonomy.

3.3. Critical Reflection

Perhaps the most ubiquitous condition of global autonomy is that of critical reflection [Frankfurt 1971; Taylor 1982; Dennett 1984; Dworkin 1988; Friedman 2003; Westlund 2003]. The basic thought behind the condition is something like the following. In common with the other animals, we are subject to various first-order motivations. Yet, unlike other animals, we have the ability to take a step back from these motivations and reflect critically and rationally upon them. We have the ability to discriminate amongst them, to endorse some and to revise or to reject others, and to seek to change ourselves in accordance with these assessments. This distinctive human capacity is, in turn, taken to be central to our status as autonomous agents. As Gerald Dworkin writes [1988: 20]:

Autonomy is conceived of as a second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes, and so forth and the capacity to accept or attempt to change these in light of higher-order preferences and values. By exercising such a capacity, persons define their nature, give meaning and coherence to their lives, and take responsibility for the kind of person they are.

As I shall now argue, such a capacity for critical self-reflection is also an important autonomy trait, since its possession helps to diminish the degree of diachronic control exercised over us by our parents and educators.¹¹ This is a thought in need of careful elaboration.

To begin, imagine two schools. Each aims to instil in its pupils beliefs that it takes to be true and values that it takes to be good. In addition, one of these schools—call it *Enlightened College*—also seeks to develop its pupils' critical and reflective capacities, so that they will be able to step back from their instilled beliefs and values, adopt a critical and questioning stance towards them, and, if necessary, revise or abandon them. By contrast, the other school—*Indoctrination High*—specifically prevents its pupils from developing such capacities. Its pupils' beliefs and values are instilled as rigid and uncriticizable systems, to be maintained and acted on without reflection

¹¹ Strictly speaking, the more important autonomy trait is critical reflection itself; possession of the *capacity* is an autonomy trait only in so far as it raises the likelihood of critical reflection actually occurring.

or scrutiny. (We may imagine their teachers as resembling Dickens's Thomas Gradgrind, 'a kind of cannon loaded to the muzzle with facts, and prepared to blow [his pupils] clean out of the regions of childhood at one discharge' [1854: 10]).

Imagine now two former students (call them *Enlightened* and *Indoctrinated*), each a typical graduate of each of the schools. Indoctrinated has never questioned nor seriously reflected upon anything he was taught at school. As an adult, he continues to hold dogmatically to the same beliefs and values that were instilled in him as a child. Enlightened, by contrast, has developed and revised much of what she was taught. Of course, she has in general developed and revised these attitudes simply in the light of *other* things she has been taught, but the process has been creative and unpredictable, and it has also been cumulative: in revising some attitudes, she has acquired a new standpoint from which to assess other attitudes, and so on. Unlike Indoctrinated, Enlightened is engaged in an ongoing process of intellectual evolution. Many will have an initial intuition that Enlightened is more autonomous than Indoctrinated. On the present account, the reason Enlightened is more autonomous lies in the ways in which she is, as an adult, less subject to the control of those who influenced her as a child. Specifically, there are two respects in which this is likely to be the case.

First, whereas all of Indoctrinated's instilled beliefs and values remain, many of Enlightened's have changed. In acting on these new attitudes, then, Enlightened is relatively independent of the wills of her former educators: for whereas they intended that she act on the basis of p , she now acts on the basis of p^* . Indoctrinated, by contrast, continues to act precisely as intended. Second, recall that the more probable A makes it that B do x , the greater the degree of control A exercises over B . Now suppose that in instilling an attitude that p , both sets of teachers intended that their pupils act (later, as adults) on the basis of p . In the case of Indoctrinated, instilling an attitude that p made it almost certain that, as an adult, he would now act on the basis of p . By contrast, in the case of Enlightened, instilling an attitude that p made it only somewhat probable that she would later act on the basis of p . This is because Enlightened's capacities of critical reflection make it possible that she will at some point revise or abandon this attitude. Therefore, even in the case of those of Enlightened's instilled attitudes that are as yet unrevised, the chains of control are weakened, and she is less subject to the wills of her former educators.

These points, made here in the context of a stylized example for the sake of clarity, apply not only to one's relationship with one's teachers but also to one's relationship with one's family and broader culture more generally. Above I seconded those who have criticized the idea of self-creation as a fantastical delusion. Yet the ideal of the critically reflective agent can be seen as, and has often been presented as, a more moderate and earthly response to similar underlying concerns [Feinberg 1989: 33–5; Taylor 1982; Dennett 1984: 74–100]. It is of course true that we are all embedded in cultural and social contexts, and that we can reflect upon and criticize these contexts only from standpoints themselves located within them. Nevertheless, the ability to reassess our received attitudes in new and sometimes unexpected ways,

and thus to contribute to the ongoing development of our own cultural traditions, renders us less under the immediate control of our forebears than we would be were we slavishly programmed to follow precisely in their footsteps. For instance, though we know that our children do not possess any metaphysically robust powers of self-creation, we nevertheless do not know just what they will do with the ideas and attitudes we bequeath them, what sense they will make of them, and how they will adapt and revise them. This makes them less susceptible to any attempts we might make to control or to manipulate their futures than if we denied them the materials needed to engage in such processes. Put most generally, a society that works to develop powers of critical reflection and thus encourages the ongoing evolution of its own cultural institutions is likely a society composed of more autonomous individuals than one that seeks to straitjacket its members with unrevisable dogmas.

4. Conclusion

Over recent years there has been a growing recognition of the fact that traditional, pure nonsocial accounts of autonomy require at least some sort of supplementation if they are to capture the concept's social dimensions. Heretofore, the main result of this has been the introduction of various hybrid accounts that supplement the familiar nonsocial conditions with a number of extra social conditions. Yet while many of these accounts do a good job of matching the extension of our ordinary notion of autonomy, they all too often tend towards mere lists of conditions lacking in theoretical elegance or unifying rationale. For instance, Oshana [2006: 75–96] lists seven separate requirements of global autonomy, yet her rationale for including precisely these seven requirements and no others consists predominantly in a direct appeal to our intuitions (albeit motivated by a series of 'case-studies' [49–74]). Though the resulting account is highly attractive, views such as this seem incomplete in so far as they leave us bereft of any principled and independent means of resolving our inevitable conflicts of intuition concerning the requirements of autonomy.¹²

On the proposed pure social view, by contrast, autonomy is analysed in terms of a single condition—resistance to interpersonal subjection—clearly derivable from the basic idea of self-rule. Moreover, as I have argued, this one property may in turn be realized by an agent's possession of any number of base properties, including many of those emphasized by hybrid theorists. So not only can the pure social account match many hybrid accounts in generating an appealing overall picture of autonomous agency, but it also improves on these accounts by providing them with an underlying rationale and, with it, a principled way of deciding just what is and is not a requirement of autonomy.

¹² Dworkin [1976] and Taylor [2009] are also, I believe, somewhat vulnerable to this complaint. Raz, who attempts to derive each of his requirements from the single idea of 'being the author of one's own life' [1986: 374], may be less so—though it is doubtful whether appeals to the notion of 'authorship' are in the end much more helpful than direct appeals to that of autonomy itself.

In this paper I have argued that social and legal protections, a healthy sense of self-worth, and a capacity for critical reflection are three properties of agents that tend to make them less susceptible to subjection by others, and therefore properties that contribute towards an agent's autonomy. Since they are also valuable properties potentially attainable by beings like us, they provide the basis for an attractive instantiation of personal autonomy. Yet there are no doubt still other valuable and relevant autonomy traits, including many other properties typically highlighted by proponents of hybrid accounts, that may also form part of a more elaborate and comprehensive picture of autonomy: properties such as rights of democratic participation, entitlement to basic levels of material sustenance, loving support of friends and family, a capacity for appropriate trust, rationality, knowledge, integrity, and creativity.¹³ More work is therefore required in order to articulate a fully elaborated account. In this paper I hope simply to have provided a conceptual framework within which a pure social account of personal autonomy may be developed, and to have indicated how such a project might profitably be continued.¹⁴

Birkbeck College, University of London

Received: January 2012

Revised: December 2012

References

- Benditt, T. 1977. Threats and Offers, *The Personalist* 58: 382–4.
- Benn, S. I. and W. L. Weinstein 1971. Being Free to Act, and Being a Free Man, *Mind* 80/318: 194–211.
- Benson, P. 1991. Autonomy and Oppressive Socialisation, *Social Theory and Practice* 17/3: 385–408.
- Benson, P. 1994. Free Agency and Self-Worth, *Journal of Philosophy* 91/12: 650–68.
- Christman, J. 2004. Relational Autonomy, Liberal Individualism, and the Social Constitution of Selves, *Philosophical Studies* 117/1–2: 143–64.
- Christman, J. 2009. *The Politics of Persons: Individual Autonomy and Socio-Historical Selves*, Cambridge: Cambridge University Press.
- Dennett, D. 1984. *Elbow Room: Varieties of Free Will Worth Wanting*, Cambridge, MA: MIT Press.
- Dickens, C. 1854 (1995). *Hard Times*, London: Penguin Books.
- Dworkin, G. 1976. Autonomy and Behavior Control, *Hastings Center Report* 6/1: 23–8.
- Dworkin, G. 1988. *The Theory and Practice of Autonomy*, Cambridge: Cambridge University Press.
- Feinberg, J. 1989. *The Moral Limits of the Criminal Law, Vol. III: Harm to Self*, New York: Oxford University Press.
- Frankfurt, H. G. 1971. Freedom of the Will and the Concept of a Person, *Journal of Philosophy* 68/1: 5–20.
- Frankfurt, H. G. 1999. Autonomy, Necessity, and Love, in his *Necessity, Volition, and Love*, Cambridge: Cambridge University Press: 129–41.
- Friedman, M. 2003. *Autonomy, Gender, Politics*, New York: Oxford University Press.
- Garnett, M. 2013. Taking the Self Out of Self-Rule, *Ethical Theory and Moral Practice* 16/1: 21–33.

¹³ Note that for each property thought to be important for autonomy, the pure social theorist must establish either that it is indeed an autonomy trait, or else that there are independent and principled reasons for excluding it from our account of autonomy. Take, for instance, identification with one's effective motivations. It may well turn out that, in general, it is easier to push around a disunited and conflicted agent than to push around an integrated and wholehearted one, and that identification is therefore an autonomy trait. If it does not, however, then the social theorist will need to join with others such as Dworkin [1988] and Taylor [2009] in providing principled reasons for denying that such identification is necessary for autonomy (see also, in this connection, n. 8, and Garnett [2013]).

¹⁴ An earlier version of this paper was presented at the 2011 New Scholars in Bioethics (NSIB) First Annual Symposium, and I would like to thank the participants (Robyn Bluhm, Kirstin Borgerson, Danielle Bromwich, Joseph Millum and Marika Warren) for their valuable comments. Versions were also presented at a number of seminars at Birkbeck College, and I am grateful to colleagues (especially Miranda Fricker, Keith Hossack and Susan James) as well as to our postgraduate students for some very helpful discussions. Finally, I am especially indebted to three anonymous referees for this journal for their thoughtful and excellent critical comments and suggestions.

- Govier, T. 1993. Self-Trust, Autonomy, and Self-Esteem, *Hypatia* 8/1: 99–120.
- Held, V. 1993. *Feminist Morality: Transforming Culture, Society, and Politics*, Chicago: University of Chicago Press.
- Hill, T. E. Jr., 1973. Servility and Self-Respect, *The Monist* 57/1: 87–104.
- Jaggar, A. M. 1988. *Feminist Politics and Human Nature*, Lanham, Maryland: Rowman & Littlefield Publishers.
- Korsgaard, C. 2009. *Self-Constitution: Agency, Identity, and Integrity*, Oxford: Oxford University Press.
- Meyers, D. T. 1989. *Self, Society, and Personal Choice*, New York: Columbia University Press.
- Nedelsky, J. 1989. Reconciling Autonomy: Sources, Thoughts, and Possibilities, *Yale Journal of Law and Feminism* 1/1: 7–16.
- Noggle, R. 2005. Autonomy and the Paradox of Self-Creation: Infinite Regresses, Finite Selves, and the Limits of Authenticity, in *Personal Autonomy*, ed. J. S. Taylor, Cambridge: Cambridge University Press: 87–108.
- Oshana, M. 2006. *Personal Autonomy in Society*, Aldershot: Ashgate Publishing Limited.
- Pettit, P. 1997. *Republicanism: A Theory of Freedom and Government*, Oxford: Clarendon Press.
- Raz, J. 1986. *The Morality of Freedom*, Oxford and New York: Oxford University Press.
- Richards, N. 1987. Acting Under Duress, *The Philosophical Quarterly* 37/146: 21–36.
- Sachs, D. 1981. How to Distinguish Self-Respect from Self-Esteem, *Philosophy and Public Affairs* 10/4: 346–60.
- Taylor, C. 1982. Responsibility for Self, in *Free Will*, ed. G. Watson, Oxford: Oxford University Press: 111–26.
- Taylor, J. S. 2009. *Practical Autonomy and Bioethics*, New York: Routledge.
- Westlund, A. 2003. Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?, *The Philosophical Review* 112/4: 483–523.
- Westlund, A. 2009. Rethinking Relational Autonomy, *Hypatia* 24/4: 26–49.