

PREPRINT – forthcoming in Journal of Moral Philosophy

PLEASE CITE PUBLISHED VERSION

Autonomy and Orthonomy

Tom O'Shea

grundlegung@inbox.com

***Abstract:** The ideal of personal autonomy faces a challenge from advocates of orthonomy, who think good government should displace self-government. These critics claim that autonomy is an arbitrary kind of psychological harmony and that we should instead concentrate on ensuring our motivations and deliberations are responsive to reasons. This paper recasts these objections as part of an intramural debate between approaches to autonomy that accept or reject the requirement for robust rational capacities. It argues that autonomy depends upon such responsiveness to reasons, countering objections that 'externalist' rationalist criteria strip the self from self-government.*

I

Does an autonomous act have to be a good or rational one? Contemporary theories of autonomy typically answer in the negative: requirements for correct actions, desires or thoughts are rejected. False beliefs, irrational attitudes, abject political commitments, or ethical failures do not themselves preclude autonomous agency on these standard accounts. Yet, this might seem puzzling, given that personal autonomy is so often taken to be a fundamental human ideal. If autonomy is equally compatible with good and bad actions, then its value can be hard to discern. Why set our sights on a normatively barren form of self-

governance, when the governing self is fallible and so liable to lead us astray? Advocating mere self-determination – without some further orientation towards what is true and good – looks to be especially ill-advised when the relevant self is damaged in some way, such as through being struck with mental disorder, colonised by malign ideologies, or sunk in moral degeneracy.

This dissociation of autonomy and correctness has led commentators to remark on the ‘evidently powerful contrast between the ideal of acting freely and that of acting rightly’.¹ When framed as a conflict, some have claimed that the latter ideal of right action is preeminent. For instance, Philip Pettit and Michael Smith recommend a shift away from ‘the ideal of autonomy or self-rule in favor of the ideal of “orthonomy” or right rule’.² This paper considers the challenge to autonomy from orthonomy and argues that conceptions of autonomy which lack robust normative conditions present us with a mistaken ideal. However, this should not lead us to relegate or abandon autonomy. Instead, we should adopt an orthonomous conception of autonomy which recognises the contribution that substantive rational capabilities make in enabling us to govern ourselves. On this approach, self-governance is always dependent upon the capacity for good governance.

¹ Susan Wolf, *Freedom Within Reason* (Oxford: Oxford University Press, 1990), p. 140. Cf. Gerald Dworkin’s attempts to dissolve the ‘conflict between self-determination and notions of correctness and objectivity’ in *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press, 1988), p. 12. For a more abstract account of the tensions between self-determination and objective correctness, see John McDowell, ‘Self-Determining Subjectivity and External Constraint’, *International Yearbook of German Idealism* 3 (2005), pp. 21-37.

² Phillip Pettit and Michael Smith, ‘Freedom in Belief and Desire’, *The Journal of Philosophy* 93:9 (1996), pp. 429-49, here p. 442. See also their ‘Backgrounding Desire’, *The Philosophical Review* 99:4 (1990), pp. 565-92; ‘Practical Unreason’ *Mind* 102: 407 (1993), pp. 53-79; and to a lesser extent, Michael Smith, ‘Beyond Belief and Desire: or, How to Be Orthonomous’, in N. Vincent, I. van de Poel and J. van den Hoven (eds.), *Moral Responsibility: Beyond Free Will and Determinism* (Dordrecht: Springer, 2011), pp. 53-70.

II

Conceptions of autonomy can perform multiple theoretical functions. These include setting boundaries to legitimate state power, identifying conditions for legal competence, and specifying the subject who is presupposed in the construction of principles of justice. Our primary focus will be autonomy as a norm of individual agency intended to guide thought and action. This kind of autonomy has also been classed as ‘an *ideal of character*, derived from the virtues and capacities presupposed by, and associated with, the condition of self-governance.’³ In this form, a conception of autonomy can steer not only individuals in their attempts to be self-governing, but also institutions which are seeking to foster autonomous agency, such as schools, supported living organisations, and social services.⁴

Orthonomy has been presented as a rival character ideal to autonomy which amounts to *good* governance rather than *self*-governance.⁵ Pettit and Smith distinguish narrow and broad versions of orthonomy. On the narrow version, orthonomy requires that our motivations are governed by our deliberative conclusions: ‘good government of desire is a

³ Steven Lecce, *Against Perfectionism: Defending Liberal Neutrality* (Toronto: University of Toronto Press, 2008), p. 103 (emphasis in original). Cf. Robert Young, *Personal Autonomy: Beyond Negative and Positive Liberty* (Beckenham: Croom Helm, 1986), ch.2, pp. 7-20.

⁴ Autonomy and related values have also been subject to criticism, of course. For the purposes of this article, I can only gesture in the direction of some of these critics: Wendy Brown, *States of Injury: Power and Freedom in Late Modernity* (Princeton: Princeton University Press, 1995), p. 157; Claudia Mills, ‘Choice and Circumstance’, *Ethics* 109 (1998), pp. 154-65; Onora O’Neill, ‘Autonomy: The Emperor’s New Clothes,’ *Aristotelian Society Supplementary Volume* 77:1 (2003), pp. 1-21; Charles Foster, *Choosing Life, Choosing Death: The Tyranny of Autonomy in Medical Ethics and Law* (London: Hart, 2009); Mikhail Valdman, ‘Outsourcing Self-Government’, *Ethics* 120 (2010), pp. 761-90.

⁵ The Greek root is *orthos* (correct, true, regular) in place of *autos* (self, same).

regime under which desire is faithful to the rule of deliberation'.⁶ Whereas, the broad version requires that these deliberative conclusions are, in turn, responsive to reasons. Against the rationalist and realist backdrop Pettit and Smith endorse, this broader ideal requires 'a sensitivity to rationally binding reasons or attunement with the world'.⁷

On its narrow construal, the concept of orthonomy has been used to diagnose distortions of practical reason concerning executive virtue. Pettit and Smith claim that it is not sufficient for an agent's desires to be governed by endogenous or non-alien forces – for example, possessing desires that the agent has decisively chosen or has endorsed by higher-order volitions.⁸ Instead, practical reason is said to require that an agent's values generate desires whose strengths are equivalent to the strengths that the corresponding values have in the agent's deliberations. This explicitly does not require that agents interrogate or give higher-order endorsement to every desire, but only that they avoid pathologies of desire – e.g. 'weakness of will, compulsions, and the rule of whim' – that would disturb the weighting of their desires relative to their values.⁹ Thus, as a narrow ideal of rightful governance, orthonomy is not primarily about the relations between lower-order and higher-order desires or volitions, but rather ensuring that a person is motivated to act in accordance with what they value in deliberation and the extent to which they value it.

Autonomy is found wanting insofar as it does not require actions to stem from correct motivations but merely chosen or endorsed ones. Of course, there are multiple senses in which motivations might be correct. On the narrow ideal of orthonomy, the requirement that motivations track values amounts to an executive virtue: it is an ability to ensure that what

⁶ Pettit and Smith, 'Practical Unreason', p. 77.

⁷ Pettit and Smith, 'Freedom in Belief and Desire', p. 443.

⁸ Pettit and Smith, 'Practical Unreason', pp. 76-7.

⁹ Pettit and Smith, 'Backgrounding Desire', p. 588.

moves us to action is what we care about as that manifests itself in our deliberative conclusions. We encounter the corresponding vice when, for example, cowardice leads us to desire our own safety more than undertaking the risky activity we conclude we should, or when due to adaptive-preference formation we abandon a difficult project we deem worthwhile. Arguably, this narrow kind of orthonomy does not conflict with autonomy – indeed, it is often understood as part of it.¹⁰ At the least, proponents of most kinds of autonomy should have little problem accommodating these executive virtues as a complementary set of abilities. However, widening the relevant sense of ‘correct motivation’ can sharpen the contrast into a potential conflict.

The broad ideal of orthonomy requires executive and substantive virtues: not only must motivation be responsive to deliberation but that deliberation itself must be responsive to reasons.¹¹ These normative constraints are taken to drive a wedge between orthonomy and autonomy. This is because orthonomy now involves ‘coming into line with something outside the realm of desire’, namely reasons for holding to the values that figure in deliberation.¹²

¹⁰ The dependence of freedom upon a person’s actions conforming to their values is defended in Gary Watson, ‘Free Agency’, *The Journal of Philosophy* 72:8 (1975), pp. 205-220. An executive dimension of autonomy is also found in non-philosophical discussions. For instance, in a medical context, the influential President’s Commission on Bioethics suggests a value-condition for autonomous decision-making, which Scott Kim glosses as being that ‘a patient’s competent choice must reflect or be consistent with her enduring values and goals’. This, he remarks, ‘has an intuitive appeal that seems to value the patient’s autonomy.’ Scott Kim, *Evaluation of Capacity to Consent to Treatment and Research* (Oxford: Oxford University Press, 2010), p. 29.

¹¹ In one of Pettit and Smith’s formulations: ‘To be orthonomous, as distinct from autonomous, an agent’s evaluations and desires have to be sensitive to his recognition of normative requirements: reasons that may be offered in support of evaluative claims.’ Pettit and Smith, ‘Freedom in Belief and Desire’, p. 443.

¹² Pettit and Smith, ‘Freedom in Belief and Desire’, p. 443.

Pettit and Smith take this to imply that it is ‘distinct from any sort of internal harmonization’ and thereby autonomy.¹³

The main objection from orthonomy builds upon longstanding criticisms of psycho-structural accounts of autonomy.¹⁴ Harry Frankfurt’s hierarchical approach – on which the relevant kind of freedom for a person is the ability to secure ‘conformity of his will to his second-order volitions’ such that ‘he is free to want what he wants to want’ – has been the basis of these accounts.¹⁵ Against this view, Pettit and Smith deny that the presence of controlling higher-order desires is significant, since they may be no more than hang-ups or products of conditioning. These desires only ‘distinguish one pattern of internal conflict resolution – one internal state of harmony – from a different but no less eligible pattern’.¹⁶

In other words, the formal property of being a second-order volition provides no grounds for granting it a privileged authority to control other desires or actions. For example, the simple fact that someone wants to realise their desire to become a missionary bestows that second-order volition neither with autonomy nor the normative authority that would stem from such autonomy. This is because the higher-order volition might be equally ungrounded, being a product of passing whimsy, muddled deliberation, or insidious pressure from a

¹³ Pettit and Smith, ‘Freedom in Belief and Desire’, p. 443.

¹⁴ For earlier variations on this objection, see Gary Watson, ‘Free Agency’, *Journal of Philosophy* 72 (1975), pp. 205-20; Marilyn Friedman, ‘Autonomy and the Split-Level Self’, *The Southern Journal of Philosophy* 24:1 (1986), pp. 19-35; and John Christman, ‘Autonomy and Personal History,’ *Canadian Journal of Philosophy* 21:1 (1991), pp. 1-24.

¹⁵ Harry Frankfurt, ‘Freedom of the Will and the Concept of a Person’, *Journal of Philosophy* 68:1 (1971), p. 15. On the evolution of hierarchical approaches, see James Stacey Taylor, ‘Introduction’, in J. Stacey Taylor (ed.) *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy* (Cambridge: Cambridge University Press, 2005), pp. 1-29.

¹⁶ Pettit and Smith, ‘Freedom in Belief and Desire’, p. 443.

religious community. The ability to want what one wants to want loses its cardinal importance if there is nothing special about actually realising this motivational structure through exercising this ability. By recommending that the broad ideal of orthonomy replace autonomy, the ability to harmonise orrectic states gets superseded by the ability to ensure that they are responsive to the authority of reasons.

III

We have seen how autonomy and orthonomy can be cast as competing ideals, and that this leaves popular conceptions of autonomy open to the objection that mere consonance with higher-order volition is not significant enough to rival the authority of substantively rational deliberation. Yet, opposing autonomy to orthonomy might seem strange.¹⁷ Must either self-rule or right-rule take precedence? Are they not instead on the same footing, such that self-governance and good governance are equally fundamental norms? This suspicion finds support when we consider alternative accounts of autonomy that make stringent forms of rationality into an aspect and requirement of self-governance rather than a potential

¹⁷ A strategy for dissolving the conflict between autonomy and orthonomy would be to claim that one or both of them are transcendental conditions of agency: that for an action, thought, or desire to be mine at all requires them to be autonomous and/or that to be an action, thought, or desire at all is to be orientated towards (what we take to be) the true and the good. To the extent that realising autonomy and orthonomy are non-negotiable goals of beings like ourselves, attempts to trade them off against one another would look misguided. However, this strategy threatens to collapse the normative sense in which autonomy and orthonomy are intelligible ideals of character which can be pursued more or less intently. For example, if someone admonishes a person, saying, ‘you should want to live your own life’ or ‘you are not worried enough about being out of touch with reality’, then they are holding them to standards that as ordinary functioning agents they can succeed or fail at meeting. Since it is at this normative level where the apparent conflict between autonomy and orthonomy arises, then transcendental solutions seem overhasty.

competitor to it.¹⁸ Only given certain understandings of autonomy – most obviously, dominant *internalist* models which reject extensive rationality conditions for autonomy – might it need to be subordinate to an independent ideal of orthonomy.¹⁹

The challenge to autonomy from orthonomy is transformable into an intramural disagreement about competing conceptions of autonomy. Criticism that autonomy amounts to an insignificant form of psychological harmony, indifferent to a more rudimentary demand to be receptive to reasons, becomes the objection that specific accounts of autonomy without requirements for robust rational constraint will collapse into anomie. If agency degenerates into caprice or mere haphazardness – such that individuals are subservient to whatever

¹⁸ See Paul Benson, ‘Freedom and Value’ *The Journal of Philosophy* 84 (1987), pp. 465-86 and ‘Feminist Intuitions and the Normative Substance of Autonomy’, in J. Stacey Taylor (ed.) *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy* (Cambridge: Cambridge University Press, 2005), pp. 124-42; Bernard Berofsky, *Liberation from Self: A Theory of Personal Autonomy* (Cambridge: Cambridge University Press, 1995); Diana Meyers, *Self, Society, and Personal Choice* (New York: Columbia University Press, 1989); and, despite terminological differences over the use of ‘autonomy’, Susan Wolf, *Freedom Within Reason* (Oxford: Oxford University Press, 1990) and John Martin Fischer and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge: Cambridge University Press, 1998).

¹⁹ Orthonomy-based conceptions of autonomy are not the only externalist rivals to internalist accounts. A major focus of current research on autonomy is relational accounts, which are also externalist, insofar as they take autonomy to be constituted in part by a person's social relationships. I do not consider here whether the critique from orthonomy also extends to these other externalist positions, although I suspect it will if they lack the relevant robust rational conditions upon autonomy. On relational autonomy, see Marina Oshana, ‘Personal Autonomy and Society’, *Journal of Social Philosophy* 29:1 (1998), pp. 81-102 and *Personal Autonomy in Society* (Hampshire: Ashgate, 2006); Catriona Mackenzie and Natalie Stoljar (eds.), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (New York: Oxford University Press, 2000); and perhaps most promisingly from an orthonomy-perspective: Paul Benson, ‘Feminist Intuitions and the Normative Substance of Autonomy’, in Taylor, *Personal Autonomy*, pp. 124-42.

desires they happen to find themselves with – then self-government is lacking.²⁰ An ability to respond to reasons provides access to a source of authority for guiding deliberation and action that is neither arbitrary nor inexorably alien. Thus, orthonomous forms of autonomy, which require rational competencies, recognise the capacity for good governance as a necessary condition of self-governance.²¹

Why reframe the debate in this way rather than as between advocates and critics of autonomy? One advantage is that it allows debates about the importance of freedom in human affairs to be skirted. By keeping autonomy's widely acknowledged centrality fixed, the issue of rationality can more easily come to the fore.²²

²⁰ Pettit and Smith also note the threat that caprice poses to agency: '[C]onsider the sorts of desire I form if I am capricious, now being moved by this property, now by that. We all acknowledge that spontaneity is good, but were I capricious in the manner envisaged, I would be more naturally seen as enslaved rather than spontaneous. I would be the captive of present fancy and whim, a pawn in the service of every passing mood. With caprice of this kind, just as much as with compulsion, my desires take charge of my values, if indeed I can be ascribed any values.' Pettit and Smith, 'Backgrounding Desire', p. 587.

This prohibition does not rule out every possible kind of capriciousness. Consider a figure like Evelyn Waugh's Captain Grimes, who believes that happiness consists in doing exactly what he wants when he wants. If, despite his higher-order plan, Grimes can still be said to be capricious, then this need not be in any objectionable sense: it will depend upon why he acts as he does and whether he has reasons for adopting his plan of action.

²¹ A recent brief discussion of orthonomy as a kind of autonomy can be found in Lubomira Radoilska, 'Autonomy and Ulysses Arrangements', in L. Radoilska (ed.), *Autonomy and Mental Disorder* (Oxford: Oxford University Press, 2012), pp. 257-8. It is also worth noting that Pettit and Smith originally introduce autonomy as a heterodox form of autonomy before in later work stressing its opposition to autonomy. See their 'Backgrounding Desire', pp. 585-9.

²² As will be clear in what follows, the conception of orthonomy pursued here is a *rationalist* one, which treats questions about the good in terms of what reasons there are for action and thought.

If the ideal of autonomy must incorporate both executive and substantive virtues, then the challenge from orthonomy can be dissolved. However, there has been longstanding resistance to making autonomy or cognate notions conditional upon acting in accordance with criteria of rationality or goodness. Noel Annan captures the spirit of these responses when he remarks on the importance of ‘the right to allow people to go their own way even if it is to hell’.²³ Of course, the most famous of these objections has been Isaiah Berlin’s fear that positive conceptions of liberty – featuring appeals to the interests or choices of a latent rational self which is ultimately more real than its empirical counterpart – can be used to legitimate political terror and totalitarianism in the name of freedom.²⁴ I shall leave this

²³ Noel Annan, ‘John Stuart Mill’, in H.S. Davies and G. Watson (eds.), *The English Mind* (Cambridge: Cambridge University Press, 1964) reprinted in J.B. Schneewind (ed.), *Mill: A Collection of Critical Essays* (London: Macmillan, 1968), p. 40. For a similar sentiment in a medical context, consider Demian Whiting’s comments that ‘the requirement to respect patients’ autonomy or decision making seems in large part designed to protect patients from those who might wish to impose on patients values (including judgments of what is wise or prudent) that are not necessarily possessed by the patient; it can hardly be a presupposition of that requirement then that patients have the capacity to make decisions that accord with the values held by others.’ Demian Whiting, ‘Does Decision-Making Capacity Require the Absence of Pathological Values?’ *Philosophy, Psychiatry and Psychology* 16: 4 (2009), p. 343. On political justifications for this kind of normatively neutral stance, see Andrew Mason, ‘Autonomy, Liberalism and State Neutrality,’ *The Philosophical Quarterly* 40:160 (1990), pp. 433-52.

²⁴ Isaiah Berlin, ‘Two Concepts of Liberty’, in I. Berlin, *Four Essays on Liberty* (Oxford: Oxford University Press, 1969), pp. 118-72. For rebuttals to Berlin, see Charles Taylor, ‘What’s Wrong With Negative Liberty’, in A. Ryan (ed.), *The Idea of Freedom* (Oxford: Oxford University Press, 1979), pp. 175-93 and John Christman, ‘Liberalism and Individual Positive Freedom’, *Ethics* 101:2 (1991), pp. 343-59.

much-discussed issue aside in order to concentrate on a related set of obstacles to orthonomous forms of autonomy, particularly what I shall call the ‘missing self’ argument.²⁵

IV

Autonomy requiring broad orthonomy as a necessary condition would amount to ‘an external value condition (in its demand of rationality)’ of the kind that defenders of internalist accounts of autonomy reject.²⁶ John Christman – perhaps the foremost contemporary theorist of autonomy – articulates the objection like so:

only minimal ‘internal’ conditions for rationality (like consistency of beliefs and desires) would be plausible as conditions for autonomy. For to demand more [...] would make the property of autonomy divergent from the idea of *self*-government that provides its intuitive base.²⁷

Note that only some rational requirements are rejected as incompatible with self-government.

Christman later clarifies which reasons autonomous agents need not be responsive to:

‘reasons’ in an objective sense, in which responsiveness is measured by one's ability to react to what are independently specifiable as good reasons in the situation. [...] Claiming that a person is autonomous only when she is correctly reasoning in light of objective criteria not only disqualifies all of us some of the time, it implausibly collapses the quality of autonomy into merely being a clear thinker.²⁸

²⁵ This is not only due to the voluminous treatment that has already been given to the topic of tyranny, but also because it is more distantly related to the notion of autonomy as an ideal of *character* that is under consideration.

²⁶ Christman, ‘Liberalism and Individual Positive Freedom’, p. 356.

²⁷ John Christman, ‘Autonomy and Personal History’, *Canadian Journal of Philosophy* 21:1 (1991), p. 14 (emphasis in original).

²⁸ John Christman, ‘Liberalism, Autonomy and Self-Transformation’, *Social Theory and Practice* 27:2 (2001), p. 199. The concern with the nature of the governing self also reappears, such that ‘demanding that the autonomous person respond to objective reasons, in either the actual or a possible world, is too stringent a requirement for autonomy if that is meant to remain tied to *self*-government.’ (emphasis in original)

The charges against non-minimal rational constraints are threefold: autonomy becomes unmoored from the self who is meant to be governing; its conditions are too strict, such that no-one would unfailingly avoid heteronomy; and it overintellectualises autonomy, making it no more than a virtue of thinking. Let us take these in reverse order.

Does orthonomous autonomy simply amount to clear thought? This objection takes the responsiveness to reasons condition to imply that autonomy is simply reasoning or intellection.²⁹ Autonomy, so understood, becomes the province of ‘the disengaged cognizer’ and fails to admit ‘an affective element’.³⁰ However, orthonomy conditions involve both evaluations and desires being sensitive to reasons, as we have seen. This means that cognitive awareness of reasons is insufficient – these reasons must shape and be integrated into our emotional dispositions and character rather than remain inert intellectual commitments. Nor is our access to reasons an affectively disengaged one: emotions are a vector of rational competence and not an alternative to it.³¹ Even if the relevant reasons are impersonal, then

²⁹ Even assuming this was correct, it would not, however, necessarily follow that autonomy was merely a kind of transparent thinking, since correct reasoning can be understood in weaker and stronger senses. The judgements or valuations being reasoned about are correct in the weak sense when they are justified in light of the available evidence, whereas they are correct in the stronger sense when they are true or valid. If orthonomous self-government depends upon strongly correct reasoning, then autonomy would only be fully realised not simply by clear, careful or otherwise diligent thought but by good thought that accurately reflects what relevant reasons there actually are.

³⁰ Christman, ‘Liberalism, Autonomy and Self-Transformation’, p.203.

³¹ Rationality need not be confined to explicit deliberation and reflection; ratiocination can be contrasted with a more diffuse sense of rationality as ‘a constellation of desires, drives, emotions, moods, experiences, images, thoughts, values, expectations, associations, dispositions, sensibilities, and commitments that take shape over a lifetime [which] comprise the many ways in which the agent’s psyche and its embodiment equip him to be responsive to reasons, with or without the blessing of his judgmental or reasoning self.’ Peter Railton,

feelings like concern, alienation and contentment can still prompt us to recognise and enact what independently needs to be done.³²

Does orthonomous autonomy impose excessive demands? The severity of externalist rationalist criteria is meant to leave its advocates not casting the net of autonomy widely enough, by excluding all agents some of the time and some agents most of the time. If our conception of autonomy is meant to determine the nature of the shield citizens can wield against paternalistic intervention from the state, or to specify the kind of agent that principles of justice should presuppose, then a narrow population of the autonomous may be problematic. Yet, given the task is to understand autonomy as an ideal of character, its relative onerousness should be no objection, so long as achieving or approximating it is still a live possibility for some of us. Indeed, an ambitious standard for self-governance has the virtue of discouraging complacency in individuals and institutions aiming to realise autonomy. Nor do attributions of orthonomous autonomy have to be construed in binary terms, unlike standard tests for whether a person currently has the legal capacity to make a decision for themselves about where to live or whether to accept a medical treatment. Instead, orthonomous forms of autonomy can be more or less fully realised along different dimensions, such as the extent to which they fully achieve the relevant executive and substantive virtues.

V

‘Normative Guidance’, in R. Shafer-Landau (ed.), *Oxford Studies in Metaethics: Volume 1* (Oxford: Oxford University Press, 2006), p. 24.

³² For a different and more sustained objection to over-intellectualising autonomy, which targets excessive requirements for critical self-reflection, see Robert Noggle, ‘The Public Conception of Autonomy and Critical Self-Reflection’, *The Southern Journal of Philosophy* 35 (1997), pp. 495-515.

The most serious charge facing orthonomous autonomy is that the requirement for robust responsiveness to reasons precludes government by the self. Recall what the implications of this would be. If autonomy cannot or should not include orthonomy conditions, then the conflict between autonomy and orthonomy reappears. Autonomy without an orthonomous dimension is vulnerable to charges of arbitrariness and anomie. In this case, autonomy looks to be a kind of freedom not worth having, and therefore the wrong ideal for us to aim for.

Undermining the self of self-government introduces two main dangers. The first is analytical: without being sufficiently anchored in the self then a conception of autonomy risks running together distinct aspects of agency and practical reason in ways that will muddy our understanding of both. Gerald Dworkin seems to have had these worries in mind when distancing himself from accounts on which ‘all good things go together, and so an autonomous person will (or must) be virtuous, nonneurotic, rational, and so forth.’³³ There is some leeway in how we construct and construe concepts such as autonomy, but decisions about this are not without consequence, and we may end up blunting our conceptual toolkit if we are not mindful about conflating independent psychological phenomena or virtues of character in a single notion. Conceptions of autonomy that do not have the self at their heart are likely to lead to confusion.

The second danger is normative, namely that muting the self in self-government means that we miss the value that is conferred by the fact it is *us* both governing and being governed – which is surely the whole point of vaunting autonomy in the first place. However interesting or worthy an ideal of self-government built around the merits of rationality and successful practical and theoretical activity is, it cannot fail to foreground the crucial normative contribution of the self which is doing the governing.

³³ Dworkin, *The Theory and Practice of Autonomy*, p. 161.

Introducing otherwise benign influences from outside the self can even be counterproductive in the production of value grounded in distinctively self-determined activity. Consider Richard Arneson's discussion of claims that certain kinds of state intervention in order to promote the good life can be self-defeating if the achievement becomes the state's rather than the individual's. He illustrates this through the example of trying 'to help one's son walk on his own by "helping" him in ways that guarantee that his movements will not qualify as walking on his own'.³⁴ Similarly, if it is not primarily the identity and commitments of their actual self, but rather rationality, that is to guide a person, then whatever the merits of the resulting thoughts and actions, it may seem that their origin in self-governance is not what makes them worthy, and thus that their value is not rooted in autonomy.

VI

Does orthonomous autonomy fall victim to the problem of the missing self? In considering this question, it is striking that even explicitly internalist accounts of autonomy adopt some requirements for responsiveness to reasons. We shall see that this suggests that the real debate is not about whether rationality is an integral component of self-government but about how robust the responsiveness to reasons criteria must be. Furthermore, in light of the ubiquity of these rationality requirements, orthonomous approaches to autonomy may not be significantly more vulnerable to the problem of the missing self than their rivals are.

³⁴ Richard Arneson, 'Liberal Neutrality on the Good: An Autopsy' in S. Wall and G. Klosko (eds.), *Perfectionism and Neutrality: Essays in Liberal Theory* (Oxford: Rowman and Littlefield, 2003), p. 197.

The doxastic requirement for understanding of relevant information is commonplace among internalist theories of autonomy.³⁵ This extends to cognate concepts such as liberty, where even many of those renowned for championing people's actual decisions and wishes accept the need for this kind of reasons-responsiveness. Most famously, J.S. Mill claims that it is not a real infringement of liberty to prevent a man from crossing an unsafe bridge if there is no time to warn him of the danger.³⁶

If understanding is a necessary condition of autonomy, then autonomy will require forms of epistemic rationality. Arguably, these epistemic conditions should be externalist in

³⁵ For instance, see Thomas Beauchamp, 'Who Deserves Autonomy, and Whose Autonomy Deserves Respect?' in J. Stacey Taylor (ed.) *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy* (Cambridge: Cambridge University Press, 2005), pp. 310-29, p. 314.

³⁶ Mill's rationale is that 'liberty consists in doing what one desires, and he does not desire to fall into the river.' John Stuart Mill, *On Liberty*, E. Alexander (ed.) (Ontario: Broadview Press, 1999), ch. 5, p. 146. In this example, achieving liberty imposes no constraint on what one desires. Nevertheless, an externalist criterion of liberty is introduced, requiring understanding of features of the world relevant to a person's actions. Given that the man does not desire to fall into the river, then the unsafeness of the bridge provides a *prima facie* reason not to cross it. Admittedly, once the man knows about the risk, Mill believes that the choice to cross should be his. In light of this, it may be more accurate to say that Mill's position implies that liberty is dependent upon the recognition of facts which impact upon what we may have reason to do. This is a weaker form of the requirement for reasons-responsiveness, which demands only the ability to respond to antecedent reasons rather than the exercise of this ability. Nevertheless, before the man is told of the danger, his lack of awareness of the *prima facie* reasons provided by the bridge's unsafeness makes it the case that stopping him crossing does not infringe his liberty. Thus, for Mill, the possession of liberty – not simply the permissibility of overriding it – is relative to a person's responsiveness to reasons, in the shape of their epistemic access to empirical facts.

Given Mill's totemic status in these debates, it is also worth noting that he thinks other deficits of practical rationality can justify coercive interference, even were the man to know about the bridge – most significantly, impairments in the weighing of risk, such as those induced by 'some state of excitement or absorption incompatible with the full use of the reflecting faculty'. Mill, *On Liberty*, p. 146.

nature: for, it is neither a necessary nor sufficient condition of understanding that we form beliefs that are fully consistent with our other doxastic or practical commitments. Understanding relevant information often involves a willingness to discard old beliefs in virtue of evidence, irrespective of whether these beliefs are internally consistent with each other. In the words Christman uses to characterise his externalist opponents, we might even say that understanding is “measured by one’s ability to react to what are independently specifiable as good reasons in the situation”.³⁷ If so, the problem of the missing self – the split between governing and governed self – is introduced by this rationality condition no less than it is by the parallel requirement in practical rationality. Whether or not the necessary condition of understanding must be rendered in terms of externalist rationality, it is nevertheless notable that self-government is tied to some significant form of rationality requirement.

At this point, it might be thought that a significant confusion is afoot. An objector could caution us about conflating actions based on informed decisions with autonomous actions. Grasping a significant amount of relevant information will often make any resulting autonomous actions more instrumentally valuable; but this is not the same as understanding being a necessary condition of autonomy, or so this objection would run. In this case, rationality could be much less central to self-governance than the appeal to understanding seemed to imply.

Yet, even accepting this, avowedly internalist theorists of autonomy introduce rationality requirements for other purposes. Recall that John Christman’s conception of

³⁷ Christman, ‘Liberalism, Autonomy and Self-Transformation’, p.199.

autonomy precludes deliberation that involves ‘*manifestly inconsistent* desires or beliefs’.³⁸ This is a rationality criterion whose aim is to rule out self-deception, since a self-deceiving agent cannot adequately self-govern.³⁹ Again, as with the necessary condition of understanding, there are grounds for questioning how internalist this manifest inconsistency condition is. For instance, what if someone rejects consistency as a norm of deliberation, or has a different idea of what manifest inconsistency involves than the rest of us? The self of autonomy is not, for them, one that can be fully identified with their actual believing and desiring self. This suggests that the problem of the missing self – namely, the criticism that orthonomous autonomy implies an unacceptable attenuation of *self*-government – can redound upon those who raise it. Whatever the success of these additional ‘companions in guilt’ arguments, it should now be clear that the disagreement with internalist accounts of autonomy is not over two radically different conceptions of self-government, but is instead about precisely how extensive the rationalist components of autonomy are. With this result in mind, we can now turn to face the problem of the missing self directly.

VII

How can robust orthonomous requirements escape being alien impositions that would tyrannise the actually existing self? If we were the sources of normativity in practical agency, then this could explain how the requirement for autonomous actions to be good or rational ones could be compatible with self-governance. Orthonomy would amount to a form of self-

³⁸ Christman, ‘Autonomy and Personal History’, p. 15: ‘Those preferences or beliefs that are in obvious conflict, ones which the agent could bring easily to consciousness and recognize as incompatible, are what one would label “manifestly inconsistent”.’

³⁹ This is said to be because such a self is ‘dissociated, fragmented or insufficiently transparent to itself’. Christman, ‘Autonomy and Personal History’, p. 17.

rule, since we would be responsive to reasons whose normative authority ultimately stemmed from our own selves.

However, anti-realist metanormative positions are not only deeply contested but also risk cutting against the main attractions of orthonomy. As Pettit and Smith note, the significance of orthonomy will be relative to the ‘metaphysical underpinning of values’.⁴⁰ Orthonomy’s criteria of correctness for action and deliberation are too weak if their content is determined by what we happen to value, desire, or take to be correct. The more psychologistic our metanormative position – particularly, the closer to emotivism or other forms of subjectivism – then the less likely orthonomy is to differ from mere psychological harmonisation.⁴¹ Anchoring orthonomy to the self through an appeal to anti-realism about value comes at the cost of diminishing the functions which orthonomous constraints can perform. Thus, it becomes self-defeating.

On the assumption that we are not the sources of normativity, making autonomy conditional upon orthonomy will mean that autonomous agents must be receptive to reasons which are not of their own making. Yet, such a position may seem wrongheaded or even perverse. Deference to authority which originates outside the self appears paradigmatically

⁴⁰ Pettit and Smith, ‘Freedom in Belief and Desire’, p. 442.

⁴¹ Problems will still arise for metanormative accounts which make human agents into the sources of normativity, whilst striving for robust forms of objectivity in their normative standards, such as some forms of Kantian constructivism. To the extent that these norms are objective and can put pressure on a person’s actual commitments, desires and beliefs, then the threat of the existing self being alienated from the rational standards governing them looms. Thus, for sufficiently objectivist metanormative accounts, the problem of the missing self reappears. For discussion of Kantian forms of constructivism, see John Rawls, ‘Kantian Constructivism in Moral Theory,’ *Journal of Philosophy* 77:9 (1980), pp. 515-572 and Christine Korsgaard, ‘Realism and Constructivism in Twentieth-Century Moral Philosophy,’ *Philosophy in America at the Turn of the Century: APA Centennial Supplement to The Journal of Philosophical Research* (Charlottesville, Virginia: The Philosophy Documentation Center, 2003), pp. 99-122.

heteronomous. In which case, building a requirement for such deference into the necessary conditions of autonomy invites paradox. Consequently, the task is to explain why the rationality requirements imposed by the ideal of orthonomy do not demand the cession of control over one's actions to an invasively alien force. Two main considerations demonstrate this: the rootedness of rationality within human lives and the ability of reason to achieve a harmonious relationship with other aspects of our selves.

Autonomy can be threatened by different exogenous influences upon human agency: from the direct physical coercion of a jackboot to the ribs, or the emotional manipulation of a wheedling friend, to our uncritical complacency in the face of dogma, or acquiescence to people occupying powerful social roles. To describe these as 'exogenous' influences suggests that the self has an inside and outside, while also taking an implicit stand on where the metaphorical boundaries lie with respect to bodies, affect, thought, and sociality. Orthonomy's metanormatively robust forms require responsiveness to reasons whose normative authority is not primarily dependent on things usually taken to fall inside these bounds. But whilst the authority of reasons might be exogenous, this does not mean that rationality – understood, substantively, as the capacity to recognise and respond to those reasons – cannot be endogenous.

Rationality has a plausible claim to be one of the cardinal capacities possessed by the self.⁴² Indeed, there are a whole range of activities performed by mature humans in which the

⁴² There are compelling reasons for resisting its inclusion as a *necessary* condition of human selfhood, since this would have problematic repercussions for the status of people with intellectual disabilities. In this respect, Michael Bérubé has warned against “a performance criterion for being human. And any performance criterion – independence, rationality, capacity for mutual cooperation, even capacity for mutual recognition – will leave some mother's child behind. It will create a residuum of the abject, a fraction of the human family that is to be left out of the accounting.” Michael Bérubé, 'Equality, Freedom, and/or Justice for All: A Response to Martha

discrimination between what is and what is not a reason is integral (e.g. arguing, commending, planning, and so on). Not only is the exercise of rationality vital to what most of us do much of the time – suffusing our understanding of the world and practical orientations toward it – this capacity is recognised in many people’s conceptions of themselves (not least many philosophers’ self-conceptions). In other words, the notion that being a rational animal is part of who and what we are continues to have currency.

Furthermore, the exercise of substantive rationality amounts to more than the mere consonance of actions with what the relevant reasons dictate: people must act *on* those reasons, which involves at least an implicit recognition of them *as* reasons. Notably, this kind of recognition is not subjection or obedience to another person’s will but an openness to the world. In light of these considerations, sensitivity to the normative authority of reasons is difficult to align with deference to the social authority of an institution or cowed assent when confronted with someone’s brute coercive power. Although there remains a gap between the self and the reasons it is called to recognise, the grounding of rationality in human activities and identities provides a bridging point that makes it hard for orthonomy to be understood as ineluctably heteronomous.

VIII

If rationality is both endogenous and the means by which our actions are governed, then orthonomy can still be *self*-government. In exercising this capacity – even in recognising and acting on reasons whose authority is exogenous – the actual self is operative rather than missing. However, I shall claim that orthonomy also threatens to conflict with autonomy in another fashion – or, more accurately, that the orthonomous and non-orthonomous

Nussbaum’, *Cognitive Disability and Its Challenge to Moral Philosophy*, E. Feder Kittay and L. Carlson (eds.) (Oxford: Wiley-Blackwell, 2010), p. 100.

dimensions of autonomy can come into conflict in ways that can prevent us governing ourselves.

In short, the rational demands of orthonomy can be tyrannical, subjugating rather than genuinely governing other elements of our selves. If someone's life is riven between rationality and their deeply felt commitments – between what reason dictates and what they experience as valuable and worthy in other respects – then cleaving to rationality at the expense of all else will not lead to autonomy. The fractured and fragmented self, whose faculties are discordant rather than working in unison, only governs itself in an attenuated form.

Friedrich Schiller provides conceptual resources for thinking through this threat from the rational demands of orthonomy.⁴³ The traditional philosophical concern about the relation of reason to other aspects of the self has been elements like desire resisting the control of rationality.⁴⁴ Yet, Schiller notes twin dangers: not only 'sensuousness, ruling over reason' but also 'reason, ruling over sensuousness'.⁴⁵ The first is compared to ochlocracy (mob rule) in which the rightful authority of reason is usurped, whereas the second is compared to

⁴³ Friedrich Schiller, 'Über Ammut und Würde', trans. George Gregory as 'On Grace and Dignity' in *Friedrich Schiller: Poet of Freedom* vol. 2, The Schiller Institute (ed.) (New York: Farrar, Straus & Giroux, 1976), pp. 337-93. For discussion of Schiller's position, see Frederick Beiser, *Schiller as Philosopher: A Re-Examination* (Oxford: Oxford University Press, 2005), pp. 106-10 and Anne Margaret Baxley, *Kant's Theory of Virtue: The Value of Autocracy* (Cambridge: Cambridge University Press, 2010), ch. 3.

⁴⁴ Hume wryly notes the preponderance of these appeals to reason: "Nothing is more usual in philosophy, and even in common life, than to talk of the combat of passion and reason, to give the preference to reason, and to assert that men are only so far virtuous as they conform themselves to its dictates. Every rational creature, 'tis said, is oblig'd to regulate his actions by reason; and if any other motive or principle challenge the direction of his conduct, he ought to oppose it, 'till it be entirely subdu'd, or at last brought to a conformity with that superior principle." David Hume, *A Treatise of Human Nature*, bk. 2 pt. 3 sc. 3.

⁴⁵ Schiller, 'On Grace and Dignity', p. 363.

monarchy in which reason simply dominates recalcitrant impulses. In the latter case, when reason merely suppresses other parts of the self, ‘man here cannot act with the totality of his harmonising nature’.⁴⁶ Schiller proposes that the highest expression of freedom is neither the subjugation of the rational by the non-rational, nor the converse, but a harmonious relation between them. In practice, this requires ‘aesthetic’ education, through which we cultivate and reform those parts of ourselves such as the passions which have often been dismissed as non-rational. Whilst reason is sovereign – so, not just any psychological harmonisation will do – it must nevertheless avoid being tyrannical if human freedom is to be fully actualised.

If rationality is not adequately integrated into our psychologies, then acting on the demands of orthonomy will lead people to be at war with themselves, forever having to struggle and strain to act as they recognise they ought. Their efforts to do so may be admirable and can even seem heroic in certain lights – they display a trait that Schiller calls *dignity*, similar to what Aristotelians call *continence*. But the divided self will only ever be self-governing in a limited fashion, and its alienated state is not a paradigmatically autonomous one.

An advocate of a hyper-orthonomous conception of autonomy – which does not recognise any potential conflict between orthonomy and other requirements for self-government – may remain unmoved by these considerations. After all, why should we be concerned about brushing aside non-rational dimensions of ourselves, when *ex hypothesi* they are not reason-giving? This response comes at the high price of having to tolerate a conception of autonomy compatible with considerable self-alienation. The problem with achieving orthonomy only at the expense of suppressing contrary impulses is that this dismembers the self. For instance, even if, laudably, we manage to overcome deep-seated laziness, anxiety, or cowardice to do what we know we should, then this still means

⁴⁶ Schiller, ‘On Grace and Dignity’, p. 373.

temporarily hiving off some of our affects and habits, despite them being integral to who and what we are. However necessary, this looks more like shutting down parts of the self than comprehensively governing them.

This is emphatically not to imply that we should seek some sort of balance or trade-off between orthonomous and non-orthonomous dimensions of autonomy. As for Schiller, reason is sovereign; and on the broad understanding of it which we have been working with, it is almost tautological to say that we ought to do what reason commands. Instead, the aim is to ensure reason is embedded much deeper within ourselves than as a presence in mere ratiocination and executive control alone. This rather abstract formulation amounts to undertaking activities of self-formation that ultimately make reason's task easier: ethical upbringing, critical education, social reinforcement of judicious behaviour, cultivating felicitous intellectual, orectic and volitional proclivities, constructing environments that foster good habits, and so on.

Of course, despite reason being capable of becoming an authentic influence, in tune with the rest of our identities and emotional lives, the struggle to integrate our rational capacities within our wider set of desires and commitments will still present us with continual challenges as individuals, and in particular cases such reconciliation will still be very difficult to achieve. Morality can still chafe; the logic of an argument can be incontrovertible and yet leave us cold; we can still be pained to believe our eyes. Yet, although we might end up with a fractious and warring psyche, where reason, will and affect are at odds, there is no reason in principle why externalist rationality prevents any such conflicts from being resolved. While reason so understood can be an inauthentic influence upon us, it need not be, and so orthonomy and self-government can be compatible.

Conclusion

Orthonomy raises the spectre of the arbitrariness of autonomy; but this is best construed as an objection that plays within debates between competing accounts of the role of rationality in autonomy. This paper has argued that arguments from orthonomy tell against internalist conceptions of autonomy which fail to include sufficiently robust conditions for motivation and deliberation to be responsive to reasons. Attempts to incorporate orthonomy as an aspect of autonomy do not fall victim to the objection that externalist criteria of rationality neglect the self of self-government. However, rationality and autonomy must be reconciled in practice and not merely in theory, such that there is real psychological work to do to overcome potentially alienating conflicts between who we actually are and who we know we ought to be.