



Humphries, James Hume (2017) *Autonomy, authority, and anarchy*.
PhD thesis.

<http://theses.gla.ac.uk/8020/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten:Theses
<http://theses.gla.ac.uk/>
theses@gl.a.ac.uk

Autonomy, Authority, and Anarchy

James Hume Humphries

M.A. (Hons), M.Litt.

*Submitted in fulfilment of the requirements for the Degree of Doctor of
Philosophy*

School of Humanities

College of Arts

University of Glasgow

September 2016

Abstract

The problem of the ‘mountain man’, the caricature of self-sufficiency and individualism, is not a new one for autonomy theorists. It seems plausible that there is genuine value in self-direction according to one’s deeply-held principles. If autonomy involves something like this, then anyone concerned with autonomy as a social rather than individualistic phenomenon must explain what (if anything) the mountain man gets wrong when he denies that his autonomy admits of being placed under obligations to others. In particular, the mountain man challenges autonomy-minded social anarchists: if his denial of legitimate non-voluntary obligations is correct, then it is not just the state we should reject, but *any* organising body with coercive powers. This may be consistent with individualist anarchism or right-libertarianism, but it sits ill with the social anarchist intuition that we can have genuine political obligations (albeit not to the state).

My thesis addresses this problem in three stages. First, I argue for a functional analysis of authority and autonomy: the concepts are not pre-existing “immovable objects”, but rather are defined by the role that they are intended to play in our discourse. I suggest that we need a concept of political or institutional authority in order to resolve co-ordination problems and pursue collaborative social goods, and a concept of autonomy to explain when and why self-direction is valuable.

Second, I defend a social-relational conception of autonomy. The autonomous agent is powerful and authoritative, where this power and authority is in large part constituted, rather than merely affected, by the social structures and relations that we stand in. We are powerful and authoritative (and thus autonomous), I argue, when we stand in relations of non-domination: we are not vulnerable to arbitrary interference in our lives, and this non-vulnerability is defended in virtue of recognition respect for us as agents. There are two important implications of this account: that autonomy comes with a built-in equality condition whereby everybody’s autonomy is threatened if anybody’s is, and that there is no principled distinction to be drawn between ‘personal’ and ‘political’ autonomy.

In the last three chapters, I suggest an autonomy-justified conception of authority. I argue for autonomy as the crucial collaborative good which authoritative institutions help us to pursue, and suggest that such institutions may legitimately claim authority if they act or effect actions in ways which are likely to promote or defend autonomy-constituting relations, and act or effect actions in ways consistent with maximal equal autonomy. Finally, I return to the anarchist argument, showing that while my accounts of autonomy and authority give us a plausible picture of how autonomy is compatible with genuinely authoritative institutions, this picture still has no room for the state.

Table of Contents

<i>Acknowledgements</i>	<i>iii</i>
1. Autonomy and Authority	1
2. Conceptual and Methodological Challenges	34
3. The Structure of Autonomy	75
4. Autonomous Relations, the Personal and the Political	121
5. Authority (I)	167
6. Authority (II)	214
7. Summary and Concluding Remarks	255
<i>Bibliography</i>	275

Acknowledgements

With great thanks to everyone in the Philosophy Department at the University of Glasgow for their help over the course of the last few years. In particular, thanks to Michael Brady for his incisive comments and his assistance in structuring and clarifying the final drafts; and, most especially, to Ben Colburn, for his unfailing support in far too many ways for me to list without this sentence becoming baroque.

From 2012 to 2015 I was supported by a scholarship from the Carnegie Trust for the Universities of Scotland, to whom I am very grateful.

I'd like to thank various long-suffering others for their critical analyses, editorial recommendations or general life assistance: Sam Beaton, Jen Corns, Rob Cowan, Oli Edwards, Rose Hen, Pat Kaczmarek, Indra Mangule, Chris Mills, Steph Rennick, Joanna Samson and Malcolm Wilson have all in their own ways been crucial, and I am indebted to them.

Finally, all my love to Helen and Marc, for teaching me to say what I think and to think about what I say; to Joanna, for her kindness and her conviction; and to Kira, who is already as suspicious of commands as her father is - and considerably better at resisting them.

Autonomy and Authority: The Problem of the Mountain Man

Introduction

In this thesis, I argue for a social-relational account of autonomy and a conception of authority - the autonomy-justified conception - on which authority is defined functionally according to the goal of maximal equal autonomy in that social-relational sense. The motivating force for the argument comes from two familiar intuitions: that there is great value to the autonomous life, and that living such a life need not imply an anti-social or profoundly individualistic existence. The former is perhaps more often endorsed than the latter, but that we can be autonomous while (or through) living a life richly laden with obligations to others is implicitly suggested by anybody who holds that friendship, family and community can be important parts of an autonomous life.¹

These intuitions are challenged by what I call the mountain man problem, the hypothetical (and, in fact, actual) interlocutor who asserts that his autonomy is simply inconsistent with any political authority or obligation. Although a general problem for political authority, the mountain man is a particular challenge to democratic authority justified through autonomy, and to the view that although states must fail to respect autonomy, not all political authority is inconsistent with autonomy.

This challenge dates at least as far back as RP Wolff, although there are parallel arguments to be found in, for example, Marx's attack on proto-anarchists and anarchists and, from the other end of the spectrum, in Stirner's thorough-going egoism.² The central claim of my thesis is that although we ought to buy at least part of Wolff's conclusion - autonomy really is inconsistent with the state - being able to consistently reject state authority without falling into an implausibly

1 See, for example, Friedman (2003), Oshana (2006), and Meyers (2005).

2 See Wolff (1998), Stirner (2009), and Marx (2008). More recently, see Miller (1984) for a useful critical overview of the individualist versus communist discourse.

individualistic account of autonomy requires that we re-assess our conceptions of both autonomy and authority. In other words, autonomy-minded anarchists should reject state authority but not political authority, because political authority is a requirement for the kind of social and egalitarian autonomy (rather than autonomy as unrestricted choosing) which anarchists aim at.³

There is a great deal of preparatory work to be done first, however. Two questions arise immediately. What is it for autonomy and authority to be in tension in the first place? What are the concepts being used here such that they seem incompatible? In this chapter, I suggest a plausible initial picture of autonomy and authority, and show how the problems raised by the mountain man are not easily solved by democratic procedures.

Although, obviously, there will be detailed examination of both notions later in the thesis, it will be helpful to make sure that we're all on more or less the same page from the start.

Autonomy is, in its etymologically strictest and least interesting (for our purposes) form, self-legislation.⁴ We are autonomous just insofar as what we do follows from what we want to do (or think we ought to do, think would best fit with our goals, and so on - I leave this very vague for now). The autonomous agent isn't a wanton choice-making machine, nor are they the psychological puppets of someone else; rather, their actions are determined by some reasonably consistent and consistently endorsed set of principles which the agent holds. Thus, I autonomously eat a cheese sandwich if I decide that eating the sandwich is the best way for me to pursue my goal of having lunch, I eat less autonomously if I am only eating the sandwich because my mother has pressured me into it; and if I have some unusual dairy addiction or have been hypnotised into believing that I like cheese, my sandwich-eating is not autonomous at all. Such a bare-bones picture leaves much to be explained - for a start, we might

3 Throughout this thesis, I use "anarchist/anarchists" to refer exclusively to the communist anarchist position unless specifically noted otherwise - whether or not Stirner's individualism, for example, has any conceptual or historical claim on the word, it is not one of the family of political theories being picked out when I refer to anarchist thought.

4 See Christman (1988) for a useful survey of various conceptions of autonomy.

want to ask questions about *how* and *why* I come to believe that eating the sandwich best serves my lunch-based principles - but it does, I believe, capture the central intuition that most of us have about autonomy: it involves directing our lives as best we can to fit with our preferences and our beliefs about what is valuable. Minimally, then, autonomy requires that we be powerful and authoritative over our lives; powerful so that we can actually exercise our self-direction, and authoritative because this direction is specifically *ours* to decide.⁵⁶

Authority, on the other hand, can be thought of as the right to legislate (broadly understood) for others: I am authoritative over you, or over something, insofar as I can demand that you do or refrain from doing something and have this demand take on a moral nature.⁷ Although the idea of a morally required obedience does not translate well across all senses of authority it is clearly relevant in terms of political authority, or the ability of some institution to legitimately command obedience of some group or groups of people.⁸ In particular, political authority seems to have some requirement that an institutional ability to command be connected with features of the institution itself. For example, the putative political authority of the British government is taken to be correlated with its following certain democratic procedures, being constituted in such-and-such a way, so that it may *not* demand obedience without satisfying the correlative requirements (and/or that it only has authority in a realm delineated by these requirements).

This first pass at autonomy and authority gives us the following very bare-bones definitions. Autonomy is self-direction according to one's goals and values; the autonomous agent's life is more or less under her control. Authority is the ability

5 I return to these requirements, and examine what they entail, in Chapter 4.

6 Although each of them gives a different conception of autonomy, this very broad and rough intuition is shared by Christman (1989), Colburn (2010), Friedman (2003), Oshana (2006) and Raz (1986) among others. Even those who - as I do, following Oshana - argue that autonomy is largely a matter of external circumstances more complicated than simply the satisfaction of authentic desires are clearly concerned with *what it is to be* self-directing in the appropriate way, rather than denying that self-direction is an appropriate component of autonomy.

7 See, for example, Carr (1983), Raz (1979) and Wolff (op. cit.).

8 Whereas we 'obey' an epistemic authority by believing them or perhaps by adopting the recommended course of action, and any moral imperative is often indirect if it is present at all.

to command others, backed with moral force - when we are given an authoritative order, we ought to obey.

So far, so (relatively) uncontroversial. Although not all political theorists attach as much importance to autonomy as I do, it is a very common value in, at least, Western liberal thought. Authority is even more widely accepted as a concept which must figure in any account of political institutions - arguably, the notion of authority is the most crucial conceptual component of the state, often used as a distinguishing factor between government and a friendly society (or government and any other band of thugs). Clearly, then, any account of political institutions - at least, any autonomy-minded account - must find a principled way of incorporating both of these concepts. We want, in short, some way to reconcile the importance of autonomous self-direction with the value of collective and authoritative decision-making: some account which vindicates the intuition that autonomy is valuable, and explains why the value of autonomy (correctly understood) does not entail rejecting any and everything which would restrict our range of choices.

It is here that the mountain man emerges from his bunker, and his challenge is far more serious than our intuitions might suggest.

1. The Mountain Man

The mountain man archetype is known by many names, but whether we call him a survivalist, a militiaman, or a mountain man, his essential details remain the same. From Thoreau's *Walden* (1995) to the story of Christopher McCandless told by *Into the Wild* (1996), key themes recur time and again: independence, self-reliance, a rejection or standing outside of society. The mountain man is beholden to nobody and nobody is beholden to him. You may do just as you wish outside his patch of land, for that is your business, but you may not interfere with what he does on that land, nor may you make any demands of him. He would no doubt come to the aid of a hiker who fell and broke their leg on the

passing track - but, crucially, he would not be obliged to do any more than the bare dictates of morality demand, and any attempt to coerce him into compliance would mark an unjustifiable infringement of his freedom.

So much the worse for mountain man, we might think. His life is not one to aspire to, and his picture of freedom an unappealing portrait of rampant individualism with strong overtones of parody masculinism. I tend to agree.⁹ The problem for the autonomy theorist is this:

If autonomy consists in self-direction, then we cannot disagree with the mountain man's assertion that being forced to (for example) contribute to a Hiker Rescue Fund constitutes a restriction of his autonomy. We may - and perhaps should - press upon him the importance of solidarity, empathy, altruism and so forth, but insofar as we respect autonomy, we *must* accede to his demands to be left alone.

The problem extends further. We have so far only considered a sort of paradigm example, a specific problem whereby it looks like I can give no autonomy-minded reason for the mountain man to change his stance.¹⁰ But exactly the same difficulty strikes when we try to make the case for *any* autonomous agent being coerced into doing *anything* that does not figure in their plans for self-direction. If we can't make mountain man help the hiker, it's not obvious why we can make Mrs Brown next door stop plastering her house with racist posters, or what gives us the right to demand that healthcare providers be forced to provide healthcare regardless of their religious opposition to particular people accessing those services.

In short, it begins to look as if the problem is not one of autonomy and rugged individualists, but one of autonomy and authority: if autonomy is (roughly) self-direction and authority is (roughly) the right to over-ride that self-direction,

⁹ Although it does seem fair to observe that anyone who has not, on occasion, felt the urge to take to the backwoods and ward off governance with a shotgun must be a citizen of some undiscovered Utopia.

¹⁰ To be clear: we might still have *other* reasons, doubtless moral ones, to offer the mountain man. The problem is specifically that an attempt to generate obligations on the basis of autonomy seems to fail, and so I've focussed on that to the exclusion of other possibilities.

then one of the two has to go. We either drop authority and let countless petty fiefdoms spring up - the “thousand stinkworts blooming” approach - or we expropriate mountain man’s Appalachian hideaway and let the autonomy theorists go hang.

It is worth being clear about the implications of these approaches. If there is no such thing as a social or political institution with the right to make demands on us, then unless and until humans become near-angels motivated by altruism and benevolence the anarchist society does not look appealing, for the anarchist ‘society’ at best looks like a collection of family units or micro-states maintaining an uneasy peace. Kropotkin (1972) famously thought that in the absence of exploitation and competition we *would* be so motivated; but, although there is much about his account that is plausible, the claim that humans are by default moral and considerate agents is surely too contentious to rest a revolution on.

Equally, if we really are concerned with self-direction then it won’t do to dismiss the anarchist rejection of state authority as callow or narrowly individualist: there *should* be a damn good reason to prevent someone from living as they would like, and it is fairly obvious that the justifications we are given for the authority of actual states are not sufficient to provide such a reason. Even when the justifications are not actively misleading or misconceived, they go unmet: we *don’t* consent to state authority, the state does *not* make most people better citizens, and state governments *don’t* effectively represent their constituents. So the anarchist may with understandable derision ask what reasons apart from fear or narrow self-interest we have to obey the state.

The rest of this chapter is concerned with one type of response to this problem of autonomy and authority: various attempts to resolve it through the authority of democracy. I argue that none of the autonomy-minded justifications for democratic authority work given the prevalent conceptions of autonomy, and that if these arguments were the only resources available then the problem really would be impossible to overcome.

2. Autonomy and Democracy

I think, and shall argue, that the mountain man problem above gives us a strong motivation to be sceptical about one or other of the concepts of autonomy and authority outlined earlier, and that the criticisms which have been made of these concepts, although ultimately defeasible, deserve our attention. However, it would perhaps be dialectically unhelpful to move straight from problem cases to entertaining the thought that these problem cases are fatal without first entertaining relevant attempts to solve the problem. In particular, one family of arguments within political philosophy seems apt for just such a task: the group of arguments which purport to show that concern for autonomy ultimately justifies political obligation within democracies. That is, arguments from autonomy to democratic authority, if they are successful, will be able to reconcile autonomy with authority by showing that the latter is necessary or valuable for the former. Since this is one of the key goals of my project, it would be remiss not to sketch such arguments and show how I think they fail.

In the context of self-direction specifically within society, we might say that autonomy consists in shaping and bringing about the collective principles, rules and institutions of our community. When we take part in such procedures, we are after all making decisions about the course of our (though also others') lives: not always immediate and direct "what shall I have for lunch?" decisions, but nonetheless decisions which will, even if indirectly, shape how my life goes. When I decide which of three jobs to pursue, my decision reflects my values and beliefs, at least if I am autonomous, and there are obvious parallels between this and a political decision about, say, public funding for jobs training.

One particularly striking parallel is that the smaller our role in making or relative power to influence these decisions is, the less autonomous we are; just as my decision about whether or not to eat a cheese sandwich seems to become less and less autonomous (as my values play an increasingly small role in it) when it is the result of misinformation, threats, brainwashing or the like. Conversely, we are on the whole more autonomous the more integrated we are into the decision-making process - the subject of a tyrant is less autonomous than the

citizen of a democracy in virtue of having less power to take part in collective decisions.

A further, perhaps obvious, point to note is that autonomy in this sense is taken to be valuable. There are difficult questions to answer about the extent to which completely unmoralised autonomy is inherently valuable, but it seems reasonable to say that at least *ceteris paribus*, infringing or restricting someone's autonomy is bad. If we think that there is value in a self-directed life, then it follows that we have a reason, albeit a defeasible reason, to object to that self-direction being compromised. Part of why we regard the tyrant as a malevolent entity, after all, is precisely because they unfairly impinge upon the autonomy of those they rule. Notably, we tend to think that citizens under a dictatorship which allowed 'personal' rights such as freedom of expression or movement, but denied 'political' or 'civil' rights like voting and the right to association, would still be severely lacking in autonomy; not, or not solely, as part of a collective agent which can have or lack autonomy, but on the individual scale.¹¹ If I can have no say over some range of decisions which affect me (among others), then it seems plausible to say that I am less self-directed and by hypothesis less autonomous.

This is the point at which democratic solutions to the problem can be introduced. Democracy, for a concept which is so widely treated, remains surprisingly woolly. This fuzziness has a major effect on the plausibility of its various justifications - it would, for example, be significantly more difficult to run an autonomy-based justification for liberal representative democracy than it would to run a welfarist argument in support of the same.¹² Similarly, the difficulties posed by unanimous democratic systems are of an entirely different nature to those posed by majoritarian democracies, since questions of consent which afflict the latter do not trouble the former. For the purposes of this chapter, 'democracy' is thus taken to refer to a direct majoritarian system unless specified otherwise; that is,

¹¹ I return to this example in Sections 2.1 and 2.2.

¹² Such an argument would still be reliant on controversial epistemic and empirical claims, but would be much less conceptually tangled: all the welfarist has to show is that our well-being is, even if only just barely, better served by liberal democracy than by (say) direct democracy or a benevolent dictatorship. See Schofield (2006) and Mill (1974) for more on the Benthamite/Millian utilitarian justifications for democracy.

when I speak of democracy I am speaking of a decision-making procedure whereby every eligible citizen may vote on any and every proposal, and a simple majority is enough to pass the proposal.

In the contemporary literature at least, the demands of autonomy are consistently cited as motivation for democracy and for the concurrent democratic authority. Liberal-democratic systems are purported either to evince respect for autonomy by their very nature, or - more modestly - to be the “least-worst” form of coercive government in terms of trading off some autonomy in order to more firmly secure the remainder (and any other goods which might be in play).¹³ To use a familiar example, democracy is taken by those in the Rawlsian tradition to be the form of government which would be chosen by any rational agent standing behind the notorious veil of ignorance: in a democracy we have the best overall chance of our autonomy being on the whole respected or promoted.¹⁴ If we care about autonomy, then something like rational compulsion demands that we accept the authority of democratic institutions in at least some contexts.

Traditional arguments from autonomy to democracy can be split into two groups. The first take there to be one or more features of democracy which, as a matter of conceptual necessity, pay heed to or promote our autonomy. On these views, there is something about the democratic process which is inseparable from valuing self-direction; a feature, or features, of democracy explicable or valuable only by reference to the importance of the autonomous life. If such arguments can be made to work, there is shared conceptual ground between autonomy and democracy and the problem of the mountain man should at least be soluble. The problem, of course, is that if the arguments fail, then the autonomy theorist cannot consistently require democratic authority. If the arguments show further that democracy and autonomy have mutually-exclusive essential requirements, then not only should autonomy theorists steer clear of

13 Recent arguments from autonomy to (some form of) democracy, or autonomy-minded justifications for democracy, can be found in Gould (1988), especially chapters 6 and 8, and Cooke (2000).

14 This is a slightly crude characterisation of the Rawlsian attitude to democracy, which itself develops across Rawls' work, but it is fair to say that democracy fits into the broader Rawlsian picture as something that is instrumentally important for the goals of pursuing conceptions of the good and developing our moral sense: see, for example, Rawls (1996).

democracy, but the democrat cannot justify their system of organisation by any reference to autonomy.

The second group of arguments claim more modestly that democratic decision-making procedures are, in some or most situations, more likely to have a positive effect on autonomy than other processes. That may be because democracy is necessarily conceptually friendly to autonomy, in which case the second group of arguments would be entailed or implied by the first, or simply because it really is the worst form of government for autonomous agents apart from all the others. While the success of the contingent justifications would give us at least some reason to acquiesce to democratic authority (albeit a somewhat self-interested seeming reason), their failure seems not only to make autonomy incompatible with autonomy, but also to make this justificatory tactic inconsistent with the previous paragraph's. It is after all difficult to imagine how a system which was necessarily autonomy-friendly could consistently fail to respect, promote or instantiate autonomy better than any alternatives. By analogy, imagine two putative reasons to buy a coffee machine: one is that this particular machine has a feature or features which result in excellent coffee, and the other is that it will probably produce better coffee than I can currently make. If it does not in fact produce excellent coffee, then it might still deliver a better cup than I have now; but if it persistently and systematically outputs flavourless brownish water which is *less* enjoyable than my current drink, then we can be fairly certain that it lacks one or more features relevant to making excellent coffee.

In other words, the first group of arguments' failure does not entail the second group being unsuccessful - if we can find no necessary links there may still be contingent or pragmatic reasons for the autonomy-minded to be democrats - but a failure to find any pragmatic reasons rebounds "backwards" along the argument and means that there can be no possibility of a conceptual necessity.

2.1. *Justifying Democracy with Autonomy - Necessary Justifications*

Probably the boldest link drawn between the concepts of autonomy and democracy is Rousseau's.¹⁵ In the state he proposes, voting in a particular way and under particular conditions is constitutive of a valuable freedom which simply does not exist in pre-political society or when we are cut off from others.¹⁶ On an account of this sort, democracy does not just instantiate autonomy. Rather, partaking in it constitutes autonomy in a way that partaking in other decision-making procedures doesn't: we collectively decide upon the principles and policies which obtain in our society, and in so doing we realise a particular kind of agency, namely the agency of autonomous moral actors.¹⁷ If these principles are not the kinds of directives that we can legitimately demand compliance with, they are empty statements of desire or intent rather than any kind of effective self-determination; consequently a need for autonomy within society brings with it an obligation to obey certain kinds of collective decisions.

Although he speaks of freedom rather than autonomy, and is in any case positing a kind of freedom of the 'real' or rational self which would sit ill with modern accounts of personal-political autonomy, Rousseau's account gives us a useful paradigm of the first conceptual justification for democracy from autonomy. Hereafter I refer to it as the "constitution" justification: the democratic process is a process of self-direction both within the community, and as part of it - which is to say, we can play a role in the direction of the community itself. Just as being able to choose between jobs or houses allows an agent to pursue their goals and their self-conception (that is, the "kind of person" they want to be, or the values and traits they regard as important to their identity), so being able to effect change with regards to the society they live in allows this self-direction to extend beyond a notional private realm. The constitution justification, further, appears to give us another reason to reject the mountain man as the correct

15 E.g. Rousseau (1968), Book 1 Ch. 8.

16 Space precludes a full rehearsal of the preconditions Rousseau suggests, but it is worth mentioning the most important: that everyone votes with the general good, as opposed to the particular good of some interest group, or simple self-interest, in mind. There is an interesting consequence of this which overlaps with some of the problems raised later: if we do not vote for the general good, we are not engaged in the kind of activity that democracy should aim at, and thus the justification for democracy fails at the conceptual level. See Rousseau, Book 2 Ch. 3.

17 Ibid.

picture of the autonomous agent: there are some kinds of autonomy which we just can't have without being embedded or engaged with society, and autonomy within society requires, as sketched above, obedience to political authority.

The second conceptually-necessary justification is connected to but distinct from the constitution justification. That claim posits that to live in a democratic society is part of what it means to be fully or maximally autonomous, while this second approach makes the more modest claim that democracy is necessarily respectful of autonomy or of the values which we take to ground the value of autonomy (and as such, is hereafter referred to as the "respect" justification).

One way to think of this is that democracy need not *instantiate* self-direction in order to respect it. Take the example the citizen of a democracy who leads a relatively autonomous life insofar as they have their own identity, sets of values and so on, but plays no overt political role in their community. They are not politically active, do not vote, perhaps they even make the occasional assertion that they "don't do politics". Assuming the paradoxical benign dictator who rules in an enlightened and far-sighted way to further the interests of their subjects, such an 'apolitical' agent would plausibly be just as self-directed in the dictatorship as they would in a democratic system. But should they have some epiphanic experience and attempt to become politically active, they would very quickly be confronted with the limits of their self-direction.

While the democracy may not have given the agent any means of self-direction they made use of, it *did* respect that self-direction and afforded it the appropriate opportunities. Indeed, part of what it is to be a democracy simply is to afford value to self-direction.¹⁸ The autocracy, however many kinds of self-direction it may tolerate, has no basic commitment to the principles which motivate our valuing of such self-direction. Thus, one could argue that valuing autonomy justifies democratic decision-making on the basis that only democracies are necessarily respectful of autonomy by dint of their motivating values.

¹⁸ Albeit perhaps indirectly or instrumentally - a utilitarian democrat might simply hold self-direction to be good because it tends to produce the most utility or satisfy the most preferences. Nevertheless, they are still plausibly committed to valuing self-direction.

There is a third option for autonomy-minded necessary justifications of democracy; it is perhaps controversial, but, I think, interesting enough to warrant consideration. The proposal is that if we are to value autonomy inter-subjectively rather than merely egoistically valuing our own autonomy, we must value it fairly: I cannot, on the basis of a moral requirement for agential respect, demand that my self-direction or self-conception take priority over anyone else's (certainly not without some strong independent argument).¹⁹ Thus, if we find ourselves in a situation where self-directions clash, we must have some appropriate method for resolving this conflict which pays due heed to both (or all) parties' autonomy - and democracy, offering as it does the same input to all relevant agents, seems like a strong candidate.²⁰

This should not and cannot be taken to mean that the condition requires everyone to be getting their own way all the time (or no-one to get their own way any of the time, although that if anything seems less likely than its opposite), but rather that each person's need or desire for self-direction should, *ceteris paribus*, be taken equally seriously. When it comes to voting, social or economic disadvantage should not negatively affect the likelihood of achieving one's preferred result, precisely because of the normative if not practical irrelevance of wealth or personal power to the value of being able to self-direct.²¹ That is to say, although economic status might well bear on how much self-directing one is able to do - if I can't afford to go swimming every day, my chances of becoming an Olympic swimmer are pretty low - there is no reason to think that it should bear on how valuable that self-direction is, either subjectively or objectively. Prioritising the self-direction of the rich cannot be justified by arguing that rich folks' self-direction is more valuable than everyone

19 In fact, it strikes me that purely egoistic self-direction is more properly thought of as autarchy than autonomy - Raz (1986, pp379-382 and 417-419) seems to argue something similar - but the distinction is not crucial for present purposes. More importantly, the independent support for prioritising one self-conception over another cannot be found in some inequality of rational capacities - as has been plausibly argued, once an agent gets beyond a basic level of competence, we do not assign them gradated levels of agential respect (and shouldn't attempt to try). This argument is made in Carter (2011, pp538-571), and I argue for the equal autonomy requirement elsewhere in the thesis.

20 Harrison (1993) suggests that the autonomy justification for democracy works something like this.

21 This condition in turn suggests that there must be fairly stringent equality *outside* the ballot box, an implication I return to later in the thesis and which is also discussed by, for example, Dworkin (2002, pp194-198)).

else's. The same is true for the other dimensions of advantage or disadvantage which affect the opportunities one may in fact have to be self-directing, and yet still don't alter one's basic status with regards to that value.

This thought owes much to Christiano's argument for democracy from equality, but it is not quite the same.²² Christiano claims that equality is valuable and democracy is thus valuable insofar as it equalises political resources; although this claim strikes me as fairly plausible on its own merits, my argument is that respect for autonomy incorporates an equality condition whereby respect for any agent *qua* agent requires similar respect for all agents *qua* agents.²³ Democracy, then, is justifiable in part because it satisfies this condition.²⁴

I have given three potential conceptual justifications for democracy from considerations of autonomy: the constitutive claim that democracy instantiates autonomy through expanding the realm of self-direction, the respect claim that it evinces respect for autonomy by virtue of its motivating values, and the equality claim that autonomy within society requires some kind of equality between agents and their competing claims which democracy is structurally suited to provide. If these justifications go through, we have a set of reasons which amount to showing that being motivated by a commitment to autonomy requires that we have and use a notion of political authority. Albeit for somewhat different conceptual reasons, all three conceptual claims give a version of the argument that there is some necessary connection between autonomy and authority. The extent to which any of these claims can be defended is examined in Section 3; for now, I turn to laying out the contingent autonomy-based justifications for democracy.

²² Christiano (1993).

²³ By analogy: if I find some trait or skill - say, the ability to balance a spoon on one's nose - admirable, then I must offer equal admiration to every agent insofar as they are equally able to a spoon on their nose. While I may *overall* have variable levels of admiration for people who have the same skill at spoon-balancing - perhaps because person A entertains orphans with their spoon, but person B (somehow) uses it to steal from pensioners - I cannot argue that person B's *ability* is any less admirable than person A's, and if there were to be some social benefit due to all and only people who did unusual things with cutlery, A and B should be equally likely to receive this benefit.

²⁴ There are also some similarities between my version of the equality condition and Singer's argument from compromise - see Singer (1973), especially pp30-37. Unlike Singer, though, my concern is not whether fair compromise increases utility, but with whether it respects the normative basis of autonomy.

2.2 Justifying Democracy with Autonomy - Contingent Justifications

Compared to the necessity justifications above, the two contingent reasons I now turn to to support the claim that autonomy justifies democracy are relatively straightforward. They share a normative basis of respect for (or promotion of) self-direction with the preceding accounts, which is to say that our viewing them as justification for democracy is dependent upon viewing the maximisation of self-direction as, at least *prima facie*, a good thing. The first of these contingent arguments is what we might call the “maximisation” model, whereby democracy’s appeal depends upon a particular means-end calculation with regards to our goals. The second approach suggests that democracy allows us to develop valuable skills of self-direction which would either be less developed, or entirely undeveloped, in non-democratic (or non-social) circumstances.

The maximisation justification, that democracies are or are likely to be the best way to maximise our autonomy given particular conditions, can be thought of as pseudo-Hobbesian insofar as it seems to arise from precisely the same kind of thinking which characterises non-state societies as nasty and brutish.²⁵ Given absolute power to effect our wishes, as a Dr Manhattan type, we would have no need to fear or even pay heed to others in our society - our self-direction could not be influenced or limited by potential malefactors.²⁶ If we were these eminently autarchic agents, democracy would serve no purpose and indeed might even prove to be a limitation on our self-direction. The flipside of this, of course, is that we have no reason to believe that we are (or, speaking in the hypothetical, will be) these kinds of agents. Moreover, if we were to exist within the same community as such an agent, or within the same community as a powerful group of like-minded agents, then at the very least our self-direction would be under threat. Absent our being nigh-omnipotent, a Hobbesian “warre” where might makes right would thus be contra-indicated by a desire for robust self-direction.

²⁵ Hobbes (1651).

²⁶ Dr Manhattan, a figure in Alan Moore’s *Watchmen* graphic novel, is (or was) a scientist who through the usual comic-book convention of outstanding foolishness and curiously beneficial irradiation becomes an immortal, indestructible energy being with the ability to manipulate matter at the sub-atomic level.

Take this toy example: I have a burning passion for eating as much cake as possible, and am faced with two possible ways to get it. One option is to remain in a Hobbesian state of nature where I could obtain as much cake as possible by force or manipulation, but would be vulnerable to others trying the same tactic on me. The other possibility is that I agree to abide by a system of rules which, while it would prevent me from taking certain paths to cake (I could not murder, extort, etc.) would also prevent others from murdering me in their pursuit of a Victoria sponge.

Weighing these options up, it would take a fairly serious fascination with confectionery to decide that the possibility of increased cake outweighs the risk of death. The sensible thing to do, then, is to buy into some set of rulers so that having the best chance of achieving our goals for the smallest chance of being oppressed (or murdered, enslaved, and so on) is not dependent upon our being a superhuman or demigod.

This is where democracy is brought into the picture. If I vote for some relevant proposal and it wins, I have effected self-direction without having to climb over bodies to do it; if, on the other hand, the vote is lost, then while my desire has been frustrated I am not (or at least, am unlikely to be) in any imminent danger of death or imprisonment because of my pursuit of this desire. For example: in a vote over whether to spend public money on this or that proposal, the worst probable outcome is that your proposal loses. Your desires go unsatisfied, and perhaps your self-direction (even self-conception) is negatively affected - not a great outcome, but nevertheless better than the posited outcomes of trying and failing to secure your aims through sheer force of will or arms.²⁷

In essence, then, the maximisation argument holds that autonomy contingently justifies democratic authority in virtue of democracy having the smallest spread to cover - we accept the occasional frustration of our wishes in exchange for the guarantee that the most serious threats to our autonomy are socially off-limits.

²⁷ A similar calculation is implied by the methodology of Rawls' Original Position: were we to stand behind the infamous veil of ignorance, knowing nothing about what our position in society would be, it would be irrational to choose a "success/death" world over a "success/frustration" world. See Rawls (1999).

The second, “enhancement” approach takes a different stance regarding how we should see society. Where the spectre of Hobbes warned us that the community is a constant threat to our autonomy, this account suggests instead that in a democracy we can develop new kinds of self-direction as well as polishing up those powers we already have.²⁸

On such a view, undemocratic organisations are problematic not merely because they do, or may, abrogate our autonomy, but because they undermine or fail to fully utilise the capacities which (in part or in whole, depending upon the view one takes) constitute this autonomy.²⁹ Even if we were to occupy some benevolent dictatorship similar to the one considered in the preceding subsection, where all the self-directing that we as-it-happened wanted to do was permitted, we would be cut off from other opportunities to develop as self-directing agents: taking part in organising our community, coming to value new experiences and types of relations and so on.³⁰

To illustrate the point, imagine that you and nine others are trapped on a remote island, under the rule of an unseen overlord. Through various methods of eavesdropping and surveillance, this power has a fairly accurate grasp of each person’s desires, and for any one person their desires are satisfied more often than not. As it so happens, neither you nor the rest of the group are tremendously bothered about the relative lack of self-direction; you may even prefer the gilded cage of the island, without this preference being in any way disqualified on grounds of oppressive socialisation or adaptive formation.³¹

Plausibly, your life in this example is perfectly pleasant: it is of course a very precarious and (we are likely to think) submissive kind of pleasant, but these are somewhat separate criticisms. The enhancement justification can say, simply,

28 E.g. Mill (1974), Arendt, (1970), Gutmann (1995).

29 Wolff (op. cit.) takes this to be part of the Kantian argument against state authority.

30 For the moment, I remain neutral on whether these new relations *constitute* greater autonomy, or are merely conducive to it: the distinction is not particularly important for the purpose of outlining the enhancement view.

31 The question of what exactly makes a preference adaptive is a live one, and cannot be done justice as a side-issue here. However, if we use something like Jon Elster’s 1989 account of an adaptive preference as one which has been brought about by particular restrictions and would not survive the lifting of those restrictions, then we can see that the folk on the island could perfectly easily prefer life on the island to a more self-directed but arduous existence - it may be a foolish or even a craven preference, but it need not be an adaptive one.

that even if no other problems applied, this case would be less than ideal because having (for example) a conflict over jam resolved by fiat from above, even if the resolution is satisfactory, is just not as valuable as having the group work out some decision-making method or principle which achieves the same result. In a similar vein, we tend to think that children coming to think of sharing as a good (or just, or fair) thing through discussion with and paying heed to their peers is a better process than having this sharing imposed by force of parental power.

It might be thought, as Elster objects in his critique of Mill, that such arguments “put the cart before the horse” in terms of democracy’s function and value: they could be seen to suggest that democracy is valuable because it allows us to expand our capacities of reason, tolerance and the like, when surely it is precisely because we possess such capacities that democracy is to be valued.³² It is worth making clear, therefore, that the enhancement account as I have construed it is not intended to suggest that the *only* justification for democracy is this nurturing of autonomy-relevant capacities. Patently, we do not partake in democracy just for the love of the game; we might appreciate the experiences, skills and so on that we develop in pursuing political activity - and if the activity comes to naught, we might even decide that these were in the end the most valuable results - but it is a mistake to regard political participation as a sort of glorified public school debate where the outcome is ultimately less important than how clever one can look along the way.

The enhancement justification, more modestly and plausibly, says that we value democracy not *just* because it allows us to self-direct, but also because the kind of self-direction it instantiates and encourages is more extensive and involves a greater number of valuable skills/capacities than the sort of self-direction typically found in oppressive societies. A different analogy should highlight the point here. Just as we may think there are reasons to obey epistemic authorities in order to sharpen our faculties of learning and analysis, the kinds of decision-making procedures undertaken and enforced by political authorities in democracies can help to improve our capacity for (or enjoyment of) autonomy.

³² Elster (1989), pp138-39.

Close of Section 2

I have argued in this section that there are two broad categories of justificatory arguments for democracy: the conceptual arguments which claim that democracy is (perhaps uniquely) respectful or constitutive of autonomy, and the contingent justifications to the effect that autonomy is most likely to be best served by democracy. It may seem that I have crowbarred a gap between them where none ought to exist: we are more likely to achieve our goals in a democracy, one might think, if or because there is something about the democratic process which inherently tends to respect and/or promote autonomy as opposed to other collective decision-making procedures. The point of this separation is not to suggest that each of these justifications is entirely conceptually independent of the others, but rather to make it easier to examine clearly the problems which arise for the various accounts, and to see how and why they may conflict with each other in attempting to provide autonomy-minded justifications for authority.

In the next section I address this task.

3. The Problems of Democracy

Having offered five justificatory routes from autonomy to democracy in the preceding section, I now propose three examples which each appear to block some of these justifications: the familiar cases of frustrated self-direction, malevolent self-direction and persistent minorities. Taken *in toto*, they pose a serious collective challenge to any attempt to justify democracy through autonomy, and thus also to reconciling autonomy and authority.

3.1. Problems of Democracy: Frustrated Self-Direction

The most obvious and familiar instance of democracy and autonomy seemingly coming apart is a case which occurs daily in actuality, that of the lost vote.

To give a mundane example, say that I and four friends have decided to go for a drink, and are in the process of settling on a pub. Being a man of limited wealth and taste, I suggest the Dive Bar, a cheap and cheerful booze-shed, while my fine wine-loving friends prefer the Bourgeois Conceit. Neither group is prepared to back down on their selection - but neither are we prepared to manipulate, bully or coerce the other into attending our preferred choice (after all, we are supposedly friends). Fortunately, we're a far-sighted and fair-minded bunch and have prepared for this contingency, unanimously settling on a rule that any irreconcilable intra-group dispute is to be resolved by means of a simple majority vote. Naturally, their four votes outweigh my one, and so, to my presumed disgust, we set course for the Bourgeois Conceit.

The first justificatory victim of this example is the constitutive claim. If democracy is valuable or worthwhile because it extends the realm in which we can self-direct, then why should we value it when it fails to extend, or even restricts, this realm?³³ Admittedly in the pub example, the realm is fairly limited to start with and the restriction fairly minor, but it still seems that where I choose to drink is something that forms a (very small) part of my self-conception and self-direction - and in any case, scaling up the example from pub to (say) use of community resources brings the problem into much starker relief. It seems, then, odd to make the argument that democracy instantiates autonomy as a matter of conceptual necessity when we have what look like clear cases of democracy abnegating autonomy.

The obvious response here is to suggest that the voting process is itself something I value, and that although one instance of self-direction has been closed off, another equally or more valuable kind/instance has been instantiated. The strongest counter-argument of this type would be something akin to Rousseau's idea that the general will of the electorate cannot, given certain

³³ Again, the question of political obligation heaves into view. We might think that the problem for autonomy isn't so much that I don't get to go to my preferred pub, but that it would be (by hypothesis) wrong of me to repudiate the process and decision which brought this about. Nevertheless, it still seems that even if I am not morally bound by the decision there is an element of frustrated self-direction - I did, after all, want to go to my preferred pub *with my friends*, and regardless of whether it would be immoral for me not to go to *their* preferred pub, my original goal goes unfulfilled.

preconditions, err.³⁴ If the general will (of my group of friends) is for the Bourgeois Conceit, then I am simply mistaken in thinking that I wanted to go to the Dive Bar and none of my real or rational goals have been frustrated.³⁵

This is not a tack that the defender of democracy is likely to find palatable. The notion of being mistaken about my own goals (or the interests of my 'real self') such that my autonomy might best be served by dancing to someone else's tune calls up the spectre of Berlin's (1969) caricature of positive liberty, whereby more or less any interference with my self-direction can be justified in service of some higher rational self. The assertion that my opposition to going to the Bourgeois Conceit was merely the result of confusion about what I *really* want is problematic on two counts. Apart from the potential for oppression so stridently noted by Berlin, such a claim sits ill with underlying notions that the ability to pursue what we want, or even just what we *think* we want, is a valuable component of autonomy. If democracy is valuable because it allows us to self-direct across a wider scope, and this self-direction is - at least on the face of it - a valuable component of autonomy, we cannot then argue that the democratic process is to be valued because it can over-ride 'mistaken' self-direction.

If, after all, I'm supposed to prefer democracy (in part) because it allows me to achieve various goals and as such it instantiates my autonomy, then it is incoherent to respond to the frustration of these goals by claiming that democracy has instantiated my autonomy through not only blocking my preferences, but by showing that I didn't want to achieve these goals (or that I didn't even have the preferences) in the first place. As I argue later, democratic frustration of desires need not be hostile to autonomy. But the claim that democratic frustration of desires means we actually didn't have such desires to begin with is again to put democracy and autonomy in the wrong order. It's in part because we are assumed to be generally right about our own desires that democracy is floated as an autonomy-preserving decision-making process, and the Rousseau-type response gets this backwards.

34 Rousseau (op. cit.).

35 It must be presumed for this case that we are voting for where we want the group to go, rather than necessarily where we want to go, or else the whole thing is a measure of particular wills and not - at least according to Rousseau - serving the proper function of freedom (that is, autonomy) or democracy in any case.

A more plausible iteration of this argument would suggest that although I did indeed want the group to go to the Dive Bar, and that this goal has been blocked, my autonomy has not overall been reduced because there are other preferences which *have* been satisfied. The most promising candidate is some first or second-order attitude towards the process by which some (perhaps not all) of my goals, such as settling on a pub to go to with friends, are achieved. I might directly desire that we vote on it because I know this to be the procedure which will result in a choice without everyone falling out, or more likely have a higher-order wish that questions of collective direction or resource distribution be settled in a particular way.³⁶ Thus, despite being lumbered with one disfavoured outcome, I also have one or more favoured outcomes, and the autonomy loss is averted.

In the pub example, this counter-argument looks fairly strong. Presumably I *do* want my friends and I to be able to make collective decisions in a reasonably calm and fair way, and would prefer losing a vote and having to drink in a wine bar to coercing, tricking etc. my friends into drinking in a cheap pub. Unfortunately, there are multiple cases where it just seems wrong to say that we have these kinds of procedural preferences, and even more where it is patently inappropriate to suggest that our thwarted desires are outweighed by satisfying any procedural preferences we *do* have.

Imagine some distant possible world where subjects of the British state can vote on legislation and, for example, I vote against the renewal of the Trident nuclear weapons system. Assuming that I am not voting wantonly, mistakenly, and so forth, it seems most plausible to say that I simply don't give a damn how fair the process is: the thing that I desire, and the only thing that I desire, is the removal of nuclear weapons from their base at Faslane. Certainly, a successful vote against the renewal may be preferable for me - it renders unnecessary irksome actions like protest, direct action, risking arrest and so forth. But losing the vote does not mean that I will regard the renewal of Trident as "balanced out" in terms of desire-satisfaction (and as such in terms of its effect on my autonomy) by the purported fairness of the procedure which led to it. Here,

³⁶ The notion of 'procedural' preferences as a solution to voting problems is proposed and discussed at some length in Hansson (1996).

then, it seems that the procedural preference argument will support the claim that democratic processes actually form part of our leading autonomous lives iff we are imagined to have these procedural preferences with regards to every matter which is voted on - and we simply do not.

The problem also afflicts the respect justification to some extent. Holding that democracy is uniquely respectful of autonomy is in slight tension with the examples above. There is something particularly disingenuous in maintaining that you have nothing but respect for someone's self-direction whilst simultaneously frustrating it (as when states and their media assert undying respect and admiration for political protest in the abstract at the same time as they victimise and demonise actual protesters). There is however a defensive argument to be made that insofar as they offer equal formal input for each agent's self-direction, democracies *do* fulfil this condition. It can't, after all, be the case that respect for someone's decisions or their decision-making ability requires that we agree with all of them or work to bring them all about. It would not seem strange, for example, for me to say that I respect my friend's decision to study law in the hope of doing human rights work, and also to say that I think his decision is mistaken on several grounds and that I will not work to help him attain this goal.³⁷

The respect justification is thus not as immediately vulnerable to frustrated self-direction as the instantiation justification, although there is a particular kind of frustration we will arrive at in Section 3.3 which *does* look to close off respect for autonomy as a plausible motivation for democracy.

Neither contingent justification suffers too badly from single instances of frustrated self-direction: it may well still be the case that I am all-things-considered more likely to get what I want for the least cost. Indeed, assuming that I vote, lose, and suffer no bad consequences related to the way I voted rather than the fact I lost, the pseudo-Hobbesian can still assert that this is preferable to the alternative. And as per the Millian argument, I could plausibly

³⁷ Indeed, it would seem stranger if I were to devote myself to ensuring that his goals came to pass, at least presuming that we did not have the kind of intimate relationship that causes folk to be deeply invested in the personal success of the other(s).

acquire or enhance some valuable skill during the decision-making process. As I'll now argue, however, both the maximisation and enhancement justifications have more worrisome challengers in store.

3.2. Problems of Democracy: Malevolent Self-Direction

Where the previous case suggested that not getting what we want is inconsistent with using autonomy to justify democracy, the following example shows that under at least some conditions, democracy is inconsistent with somebody or some group getting exactly what they want.

Suppose that, as before, I and my friends are deliberating on how to pass the evening. No longer merely content to foist overpriced drinks upon me, the majority have now decided that only imprisoning me in stocks and throwing fruit at me will satisfy their curiously medieval desires. From my perspective, the process has thrown up the problem referred to above - if there are any situations with agents we can point to and say with confidence "they're not self-directed", being pelted with rotten tomatoes on the whim of another has got to be pushing for a podium spot - but there are further important questions to be asked about the extent to which my friends' actions cohere with the justifications given in Section 2.

That democracy respects self-direction is not just supposed to be a subjectively valuable thing, after all. While we do have good reason to value it because self-direction is (or is likely to be) valuable *to us*, justifying democracy is surely not simply a matter of asserting "well, I like it". The justification has to work for all relevantly similar agents in order to serve as a motivation for equal political organisation, and in the stocks case this doesn't happen.³⁸ If democracy can respect self-direction only insofar as we are members of a politically or socially privileged group, there is little justificatory distance between it and oligarchy or populist autocracy (which raises obvious questions about purportedly democratic modern states), and membership of the voting majority looks very like political

³⁸ A 'relevantly similar agent' can be cashed out fairly loosely at this point as simply referring to someone who is receptive to autonomy justifications (and so not, for example, concerned only with serving the will of a deity or furthering their own interests), although obviously this would need more sharpening before it could do any practical work.

privilege in this case. In other words, it seems that cases of majority-mandated immorality are permissible under democracy, and this fails the “respect for autonomy” condition for the minority. This again suggests that there is no necessary link between autonomy and democracy.

If this is so, then the idea of autonomy requiring democratic authority is vulnerable on two fronts. Either it’s simply wrong that autonomy can be reconciled with authority through democracy, or autonomy only requires us to acknowledge the authority of commands that are in accordance with our wishes.³⁹ This latter option is, fairly obviously, a cigarette paper away from being the mountain man’s position, but the former alternative is really no better. Both are useless if we are trying to reconcile autonomy and political obligation.

The response might be made that respect for autonomy requires a kind of institutional distance that is inconsistent with inspecting every act before approving it. It is unfortunate (we might concede) that the occasional malign or misconceived exercise in self-direction slips through, but if the alternative is overweening social and legal scrutiny, surely the first option displays greater respect.⁴⁰

Although this counter-argument has some intuitive appeal, it rests on a mischaracterisation. We are talking about collective self-direction in the form of laws or institutional rules, and *surely* these are precisely the kinds of things that we can, and should, give careful thought to. To give an analogy, it is absurd to expect folk to submit to (purportedly) random stop-searches just on the basis that they might, perhaps, have done something wrong or be about to do something wrong; there is no need for citizens to justify themselves against purely speculative charges of immoral behaviour. Conversely, it is not absurd at all to consider all possible reasons to implement or repeal a law requiring submission to these searches. It is the character of the proposal, not any agents, which is at issue.

39 Or perhaps just one, very wide front: autonomy can only be consistent with authority through decisions that accord with our preferences, so democratic authority has no particular autonomy legitimation *and* (insofar as no other institution can claim legitimate authority either) the mountain man is correct.

40 Carter’s (op. cit.) notion of ‘opacity respect’ suggests something like this, for example.

Thus, the objection should be read not as claiming that there is something conceptually wrong with a law which may permit malevolent acts (for example, allowing people to drive once they reach a certain age allows them much greater legally-protected freedom of movement and potentially a greater area to commit crimes across), but rather that there are some proposals or kinds of proposals that just *are* malevolent towards some group or groups. At the absolute best, such proposals seem to be respecting one group's autonomy (namely those who voted for it) while ignoring the autonomy of those on the receiving end of it. If all the respect justification amounts to is that democracy sometimes respects the autonomy of some people then it is not much of a justification at all.

We might wonder also if there is not something suspicious about the enforcement of oppression (or in the stocks case, merely immorality) being legitimated as part of the exercise of self-direction - something which would cause the constitutive justification to fail into the bargain. Admittedly, this objection is predicated on a moralised notion of autonomy and would thus face criticism from those who agree with Dworkin that the "autonomous person may be a saint or a sinner" (1989, p62). As it happens, I am far from convinced (as I argue in later chapters) by the content-neutral accounts of autonomy which are generally the basis of such criticisms. More importantly for the moment, however, the problem I now raise is not that these kinds of actions are necessarily ruled out of the realm of autonomy, but that they run afoul of the equality condition referred to in Section 2.1. That is, in order to justify democracy through respect for autonomy, it must systemically show this respect equally according to relevant similarities. In cases where immoral or oppressive proposals are mandated by the majority, this problem takes the same form as the respect problem given above, but there is an added dimension to it.

Regardless of whether or not one's own autonomy is compatible with immoral or oppressive behaviour, it is uncontroversial that there are some acts and attitudes or kinds of acts and attitudes which necessarily fail to respect *others'* autonomy: bigots cannot respect the autonomy of their targets of hatred, misogynists cannot respect the autonomy of women, and so on. There are,

therefore, some kinds of political proposals which are similarly inconsistent with respect for others' autonomy. If these proposals are passed, then patently the group or groups malignly affected by the measures have *not* had their autonomy treated equally with others. By definition, a policy which fails to respect a group's autonomy cannot be treating them equally with groups whose autonomy is respected by the policy. So, malevolent self-direction defeats the equality as well as the respect conceptual justifications.⁴¹

Curiously, given that frustrated self-direction does not pose a particular problem for contingent justifications, malevolent self-direction seems - as I'll now explain - to cut off both the maximisation and the enhancement routes.

If democracy only protects me from the baying mob so long as I am part of it (or have a bigger baying mob of my own) then, again, it looks much as if it is simply ballot-papering over the brute exercise of power. To draw out the point, imagine that I am faced with two possible scenarios, in both of which I occupy an unpopular minority of one. In the first case, the majority group decides to burn down my house, with any discussion or decision-making procedure which does occur confined to within the majority group. In the second case, the majority put the house-burning proposal to a vote I may take part in. About the best that can be said of the second case is that there is an extra step between the proposal and its enactment. I may have had some purely formal opportunity to try and prevent it that was not present in the first case, but the mere fact that democratic decision-making procedures are in place has done nothing to increase my autonomy, or even give it more stable protection than would be the case in a non-democratic society. Someone might object here that I'm applying a more demanding standard than I've been using in the frustrated self-direction case.⁴² If occasional frustrations of self-direction aren't enough to defeat the maximisation justification, why should occasional instances of malevolent self-direction do so? The reason is this: as discussed in the previous section, all

41 In fact, at this point the best bet for upholding the equality condition looks to be accepting a moralised account of autonomy and arguing from there that neither proponents nor victims of the proposal are autonomous. Even here, though, there is an obvious gap between someone who is not autonomous in the sense that fulfilling their goals would be inconsistent with maximal equal autonomy, and someone who is not autonomous because they are on the receiving end of oppression.

42 My thanks to Robert Cowan for highlighting several argumentative gaps in this section.

democracy needs to do to claim maximisation as an autonomy-minded justification is to show that democracy is better for autonomy than a Hobbesian state of nature (or an autocracy, etc). Plausibly, it does this even when our self-direction is frustrated - it's better to vote and then be frustrated than to be frustrated and then dead. But if we're talking about something like the arson case, there just *isn't* a meaningful distinction between my house being burned down by a democratically-endorsed public servant, a Hobbesian mob of amateur pyromaniacs, or an autocrat's paramilitaries. It's also not necessary for this to be, as it were, the 'standard' in a democracy. Although it's clear that this kind of case does in fact occur pretty frequently (as in, for example, the recent French bans on certain kinds of religious dress), there's a deeper problem. In particular, responding to malevolent self-direction cases by denying that they're the norm suggests that in fact, two justificatory strategies are covertly being appealed to: either authority is justified iff it actually does maximise our self-direction - in which case democratic endorsement is irrelevant and we have no reason to abide by democratic decisions which fail to maximise that self-direction - or democracy's justification is *not* actually to do with maximisation, but some other value like formal equality in decision-making.

If, conversely, I can get what I want in a democracy just by dint of membership of a powerful group - to reverse the example, if it is now me deciding whether I should burn someone's house down - then the various laudable traits which democratic decision-making is supposed to encourage have no role to play, and the enhancement justification is left to flap in the wind. Perhaps, in the discussion around the proposal, I will be struck by a particularly forceful argument and come to realise that my arsonist ways are wrong, developing a new facet of my moral character which comes to expand my autonomy. But this could equally occur if I was part of a torch-wielding mob being given an impromptu dressing-down by an onlooker. Unless the enhancement claim can produce some independent argument that democracies necessarily encourage more thoughtful (and hence, mysteriously, more autonomous) citizens than the alternative, the apparent endorsement of malevolent self-direction is a blow against this approach to justifying democracy.

3.3. *Problems of Democracy: Persistent Minorities*

The final problem I consider is that of persistent minorities, where there is some particular group (understood *de dicto*) that is bound to lose any vote, either with regards to some particular matter or, worse, any vote at all. Take, for example, a culturally-homogeneous immigrant group within some society where the dominant social mores demand short hair, as opposed to the group's preference (whether social, religious, or practical) for long hair. Imagine, further, that the short hair norm is enforced by law - it is illegal to have hair beyond a certain length - but that as a direct democracy, all laws can be challenged and voted on. Even if the group manage to challenge the law, they will be bound to lose. The fact that their group is far smaller than the opposing group means that there is absolutely no prospect of the law being overturned through the democratic process. Worse, we can stipulate without implausibility that it is not the case that the group is in the minority just because their preferences differ (as it happens) from the majority - while it is true that they are in a minority preference-wise, this is not the whole story. They are essentially excluded from collective self-direction by dint of their status as an immigrant minority; the law regarding hair length functions to coercively assimilate minority groups, and their purely formal access to democratic decision-making does nothing to change this.

The difficulties posed by the case of persistent minorities overlap to some extent with the problems raised by frustrated self-direction. The problems are not, however, co-extensive.

Firstly, the respect justification is inconsistent with persistent minority status. If democracy fails to respect my autonomy on those occasions when I am in the minority, this failure is presumably worsened by systemic inability to correct it. To put it another way, if the possibility of being in the minority once blocks the conceptual respect approach, then the certainty of being in the minority on every occasion must also do so.

More damagingly, persistent minorities also pose problems for the equality condition. While the respect angle at least has the possibility of arguing that

affording everyone the same initial input takes equal heed of the normative basis of valuing autonomy (albeit this argument is fairly weak), it is impossible to justify democracy via equality if simple membership of a particular group is a guaranteed actual or potential block to self-direction.

To clarify the distinction between this and the frustrated self-direction problem, consider the analogy of a football league. In one case, the league is set up so that each team is allocated the same amount of money for transfers, youth development and so forth each season. While, obviously, not every team can win every game, and there may be periods where one or two teams are dominant (where some particularly good group of players breaks into the first team at the same time, or the like), there is no structural barrier to any given team winning any given game - it is, as it were, a level playing field. Similarly, a single instance of frustrated self-direction only poses a problem for the equality condition if we assume that equal autonomy consists in everyone self-directing exactly as they would wish all the time. As I have argued, there is no reason to believe that the condition includes this fairly implausible requirement.

Persistent minorities, however, seem more like smaller clubs in contemporary Scottish and English leagues. Formally, there is no reason why they should not win any given game, but the structure of professional football makes it - barring deeply bizarre circumstances - impossible that they will win against more than a handful of teams.⁴³ Just by dint of not having a legacy of success, or a billionaire backer, or a whole variety of other things which should not be relevant to a contest of skill (or teamwork, or diving), some groups are ruled out. Similarly, just by dint of some socio-economic or cultural identifier, persistent minorities are structurally excluded from democracy. This is surely incompatible with the equality condition: respect for people as agents should be equal (both equally merited and equally afforded) across all agents, and membership of a persistent minority makes no odds as to whether or not one is an agent.

⁴³ Since I first wrote this, it's been somewhat overtaken by events, specifically the English Premier League being won by rank outsiders Leicester City - but the mere fact of this being such a surprise suggests that the analogy still holds up fairly well.

Drawing out the point further: the point of the equality condition is, like the presumed point of league divisions and indeed the rules of the game itself, not to ensure that everybody wins all the time, but that everyone has a fair chance of winning, or that every team's participation grants them an equal right to compete for the win. In football, the motivation for this is taken to be something like the idea that no team has a right to win just because of who they are - if, to slip into managerial jargon, the other team wants it more, they should win even if they are playing the Grand Champions of the World. League divisions are supposed to go some way towards ensuring that in any given match, the winner will be determined only by those factors relevant to competitive sport (whatever they may be). Similarly, then, the presumed autonomy motivation for democracy must say that every agent defaults to having a chance to win; that is, every agent must have the same respect shown to their proposal.⁴⁴ In the case of persistent minorities, however, the agents are *always* playing against the slope, not because they keep losing the toss, but because they had no meaningful chance to win it.

Finally and fairly obviously, the problem of persistent minorities blocks off the maximisation justification (at least from an autonomy standpoint - it may still be better in terms of, e.g., personal safety, although history suggests otherwise). After all, we cannot cogently claim that democracy gives an agent the best chance of pursuing their goals if it is structurally impossible for them to effectively pursue those goals. Again, this is distinct from instances of frustrated self-direction insofar as occasional frustrations needn't act as a defeater for the maximisation justification, whereas persistent frustration (as happens here) is just inconsistent with deploying the maximisation approach. If we want to say to someone that they should accept democratic authority because such a system maximises their chances of getting what they want, we'd better be right - and in the case of persistent minorities, we're clearly not. It also raises a side-problem for the enhancement justification: namely, whether it's even sensible to think of participation as developing valuable traits when this participation has no effective political upshot. To re-mark the Elsterian line, it seems very strange to

⁴⁴ Or, to use Singer's tongue-in-cheek but fairly accurate description of the model, we want "a fair distribution of unfairness" (op. cit., p53).

assert that a persistent minority's autonomy is well served just because voting gives them new opportunities to have a good hard think about the issues.

Close of Section 3

The problems of frustrated self-direction, malevolent self-direction and persistent minorities, taken together, look to close off every potential justification suggested in Section 2. Frustrated self-direction poses an obvious challenge for the constitutive justification and a somewhat lesser one to the respect argument. Malevolent self-direction certainly undermines the respect, equality, maximisation and enhancement justifications and has potential to block the constitutive justification; and the problem of persistent minorities applies to every justification, encompassing many of the problems of frustrated self-direction while extending them in the case of the equality, maximisation and enhancement justifications.

Furthermore, none of the counter-arguments attempted so far have gained much purchase. The likeliest-looking option, regarding the satisfaction of procedural preferences as a way to counter-balance any autonomy loss that accompanies the frustration of first-order preferences, is in essence reliant on unsound empirical claims.

Each of these problems presents a slightly different version of the same basic challenge, which was to make autonomy and authority consistent through the operation of democratic decision-making procedures. If there is no conceptual connection between autonomy and democratic authority, then there is no reason to think that concern for autonomy rationally compels us to make use of some idea of authority. If functioning democracies are not in fact the most contingently likely societies to support or protect autonomy, then there is no pragmatic reason for autonomy to motivate us towards democratic authority. In short, it seems as if the mountain man poses a far more serious problem than it may initially seem. Without being able to prove him wrong with democracy, then

his claim that autonomy and authority are irreconcilable will have to be accepted.⁴⁵

4. Where Next?

The apparent failure of democratic authority to incorporate concern for autonomy (still understood broadly as a self-directed life) leaves two obvious routes open. We might decide that the concept of autonomy is no great loss, and dispense with it; or, alternatively, reject *authority* as the misguided or misdirected value. There are compelling *prima facie* cases for both courses of action, as I outline in the next chapter. However, there is also a third possibility, that of harmonising autonomy and authority through revisiting the respective conceptions. This is the line that I will ultimately defend. In particular, as I will argue, the apparent impasse here can be resolved if we proceed via functional rather than conceptual analysis: that is, if we drop the idea that there are antecedent and uncontroversial concepts of autonomy and authority, and instead take these concepts to be most usefully delimited by reference to what purpose they're supposed to play in our political discourse.

Given that the strongest challenges to autonomy and authority are, I claim, criticisms of particular *conceptions* rather than of the concepts themselves (or the values these concepts invoke), we should therefore be motivated to perform this functional analysis rather than ejecting the baby with the bathwater.

⁴⁵ One response that I have not considered here is that autonomy may be consistent with some non-democratic authority. I assume that it's fairly obvious why the autonomy justifications already given do not militate towards such a response, and will not expound much further; there may be, although I think it's unlikely, some reasonably plausible conception of autonomy which can be made consistent with a particular kind of technocracy or benevolent dictatorship, but that conception would surely have more in common with moral than with personal-political autonomy.

2.

Conceptual and Methodological Challenges

Introduction

The previous chapter presented the caricature of the mountain man - the rugged individualist who rejects all institutional authority on the basis of personal autonomy - and the problems that frustrated self-direction, malevolent self-direction and persistent minorities pose to giving an autonomy-minded justification for his political obligation.

The problem, to restate it, is this: if we value autonomy as self-direction, it is difficult to give a reason why we should expect somebody to accept institutional obligation to do anything inconsistent with their self-direction. On the other hand, if one of the justifications for political authority that we give is its promoting or instantiating autonomy then it seems that no authority *may* claim obligations which are inconsistent with autonomy as self-direction.

I will eventually show that the conflict can be resolved. Before that task can be undertaken, though, there's some groundwork to do. First, I must defend against criticisms that (one or both of) the concepts are not worth rescuing. Second, I need to explain the methodology by which plausible conceptions of autonomy and authority can be made to cohere. In other words, the point of this chapter is to show that we can (and should) revise rather than reject our notions of autonomy and authority, and that a functional analysis gives us the best tools to do this.

The first section lays out some reasons to think that authority as a concept is not worth the hassle of keeping, before arguing that the hassle is greatly reduced if we think of the concept as being shaped by the role we want it to play rather than being an immovable object against which we bash conceptually irresistible forces. Much of the philosophical literature on authority (Wolff 1998, Simmons 1979, Carr 1983) has presumed a concept of authority which captures our pre-

theoretic intuitions and which has clear necessity and sufficiency conditions: I argue that we are better served by constructing a concept according to the philosophical or practical function the concept is supposed to perform.

The second section repeats much the same process for autonomy. Here, however, the criticisms are more numerous and the functional analysis less drawn-out, since the concept and value of autonomy have been challenged in somewhat more depth than that of (institutional) authority. I consider the arguments that autonomy as the concept has been traditionally understood is disvaluable or impractical, and show that the conceptual requirements of being powerful and authoritative over one's life capture the purpose of autonomy in our discourse and are neither objectionable nor implausible. In essence, I argue, we need a concept of autonomy that captures and explains the value of self-direction as reflecting the value of agency. In particular, the functional analysis of autonomy suggests that the concept must illustrate the notion that being able to creatively work over our surroundings and pursue our aims in accordance with our principles is a basic value for agents, and hence requires that autonomous agents be powerful and authoritative.

I conclude the chapter in Section 3 by heading off a potential distraction. I show that the concept of moral autonomy cannot play the required role: moral autonomy ultimately requires something like the ability to make suitably responsible moral decisions, and hence can be claimed without being powerful and authoritative over the direction of one's life.

1. Authority: Criticisms and Two Analyses

1.1 Against All Authority?

Given the ultimately anarchist conclusions of my thesis, it seems as well to entertain the case for a total rejection of institutional authority as early as possible. If we value self-direction, the thought goes, then we have no reason to

buy into a concept which tells us that our self-direction can be over-ridden even if we're not doing anything independently morally objectionable. If part of the concept's baggage includes making it morally objectionable not to do what an institution tells us, then so much the worse for political authority.

This is roughly the line that R.P. Wolff famously adopted with *In Defense of Anarchism*, and it is not as absurd as it may at first appear. When Wolff argues that “there is no such thing, strictly speaking, as a *command*” (1998, p15) for autonomous agents, he is not advocating a kind of nihilist or subjectivist free-for-all, nor a Stirnerian egoism where only doing exactly as you wish can be properly called autonomy.¹ In fact, Wolff's view of autonomy is a rigorously moralised one; we are autonomous insofar as we act according to rationally universalisable principles of action (in other words, insofar as we're Kantians). Political obligation would, by virtue of obedience to alien causes, involve the commissioning of heteronomous maxims, so our moral duties as autonomous agents demand that we reject any and all authority which is not moral authority.

It is true that Wolff's argument is somewhat vitiated by an apparent conflation between moral and personal-political autonomy, but as that is something I return to in the final section it can be put aside for now. The key claim can plausibly survive being disengaged from its Kantian anchor, and the key claim is that political obligation ends up being, in more or fewer words, a variation on “do x because I said so!”. This seems like a poor candidate for authority of any sort, since political authorities should give us moral reasons for, rather than simply demanding or compelling, action. If there were only prudential reasons to obey a political authority, then it would be odd to say that we had any obligation to follow their directives (although it might be prudent to do so). On the other hand, if political authority and political obligation boil down to being required to do what somebody else tells you to without any regard for your own reasons, it is difficult to see what autonomy-minded motivation there could be for accepting them. On this view, therefore, commands of the sort given us by putative political authorities are the kinds of command that the autonomous

¹ See Stirner (2009), p153.

agent, and the agent who values autonomy, have no reason to treat as *authoritative* (although they may still be compelling).

There is a notable conceptual gap here. It is only *if* political authority and political obligation are divorced from good reasons for action that they can't give us legitimate commands - so what reasons are there for thinking that this conditional obtains? Again, I will pass fairly briefly over the detail of Wolff's argument against democratic authority: in essence, and fairly similarly to some of the problems posed for democracy in the last chapter, Wolff claims that there is no appropriate connection between the commands of actually-existing states and rationally or morally forceful commands.

If we lived in a sort of quasi-Rousseauian state where we could plausibly be said to self-legislate, or if we had agreed beforehand to adopt such-and-such a decision-making procedure, then the connection *would* be there - the autonomous agent abides by their own principles, and sticks to promises and agreements - and it may be permissible for us to be coerced into fulfilling these obligations; which are, in the final analysis, self-imposed. Wolff's comment to this effect, that external compulsion under such conditions would introduce us to our "better self in the form of the state" (ibid, p23) is therefore consistent with his earlier denial of commands, for in demanding and even compelling obedience the state (or other political authorities) which operated like this would be not so much commanding us as refusing to deviate from our *own* commands.

The Odysseus example, though fairly shopworn at this point, is a useful illustration. When Odysseus requests of his crew that they tie him to the mast, he is making a commitment roughly analogous to our endorsing some policy or procedure; when they subsequently refuse to untie him despite his apparent change of heart, they are not *commanding* him to remain tied to the mast, but carrying out his own self-regarding commands. Odysseus would still, for Wolff, be autonomous despite this coercion because it is coercion that he has himself endorsed (and even required).

But actual states do not function like this: if we were to reverse the analogy and try to make Odysseus' story parallel political obligation in (say) the UK, we would probably end up with something like the following. Odysseus wakes up on a ship, with no notion of how he got there, and is told that he must either do what the crew tells him or be thrown overboard. Feeling that his odds are better on-board the ship than off it, he agrees to this ultimatum; when they pass the Sirens, he is tied to the mast and, again, the crew refuse to untie him. The Sirens' potential for interfering with preference formation put to one side, it should be obvious that in this case Odysseus' reduced freedom of action *is* an infringement of his autonomy - not because he can't get away from the mast, but because there is no point in this sequence of events where he can self-legislate; there is no sense in which his being tied to the mast, or even on board, is a state of affairs that follows from his autonomous agency.²

This, I believe, captures the key conceptual claim in Wolff's argument against the political authority of actual states (the arguments that he makes in support of the claim having been discussed in somewhat more detail in the previous chapter), but it still doesn't quite get us to the thorough-going absolutist position I suggested at the start of this section. There are after all other forms of social organisation than liberal democracy, and at least some of Wolff's supporting arguments (such as his criticism of the idea that a 'representative' can ever actually represent their constituents) are closely tied to specific kinds of putatively democratic institutions. Further, he often seems to endorse some form of direct democracy; this suggests that something exactly like the promise-keeping obligations referred to above may provide a workable - though very stringent - account of political authority whereby an institution may coerce us to comply with laws or policies that we have directly voted on, at least if we have (presumably explicitly) consented to abide by the result of the vote.

2 I leave aside the faintly facetious objection that he could have self-legislated by jumping overboard. Kant *might*, if we could paint the action as martyrdom rather than suicide, suggest that Odysseus should indeed have taken death before imprisonment; but given that the context of our discussion is one of what legitimates external coercion or institutional authority, we should go along with Hume's famous criticism (1748) of the legitimacy of our obligation in such circumstances. It seems misguided as well as incorrect to assert that you can produce political obligation through the barrel of a gun (or perhaps off the end of a plank).

There are three problems with this. First and least important (at least for our purposes) is that Wolff appears to be committed to denying that promise-keeping obligations *are* autonomous - from what he says during *In Defense of Anarchism*, doing something “because you promised!” is just as heteronomous as “because I said so!”.³

The second and third worries, however, point towards why Wolff may in fact be committed to a more anti-authoritarian position than he believes. If we ought to obey institutions just in virtue of having said that we would, it’s not obvious why these institutions have to be democratic in form. Perhaps their being democratic is something which leads us to agree in the first place, but obligation based on original consent gives no reason to think that we have stronger obligations to democracies than to any other promise-bearer. This seems, at the least, odd: even if there would be other reasons against obeying a fascist government which decisively over-ride our putative obligation, we surely want to deny that we have any *prima facie* reason to obey fascist institutions, let alone one which is on a conceptual par with reasons to obey laws in a democracy.

The second worry, then, is that Wolff seems committed to saying that *if* there can be any political obligation, it is grounded in whatever moral requirements are taken on by agreeing to bear that obligation, *and in nothing else*. In short, some kind of consent or contract theory is the only way to reconcile autonomy and authority - but in addition to the issues contract theories face with regards to when consent is genuine, whether the contract should be actual or hypothetical, etc., the problem arises that such theories give political authorities no particular institutional requirements or bounds. That is to say, if I consent to obey an unjust authority, then unless it acts in such a way as to breach whatever terms the consent has been given under, my obligation still stands however unjust the institution is.⁴ Even if we were to accept such a view

3 “It can be replied that my obligation rests upon my *promise* to obey, and that may in fact be true. But insofar as a promise of that sort is the sole ground of my duty to obey, I can no longer be said to be autonomous”. Wolff (op. cit., p29).

4 There may still be countervailing reasons to disobey (“on the one hand I ought to keep my promises, but on the other hand I ought not to support injustice”, and so on), but this is a somewhat different thing from the obligation itself being sensitive to institutional features. Again, such sensitivity *could be* built into the terms that we consent to - this seems to be

of authority, we would be left with the problem that this *cannot* be consistent with authority's being justified by autonomy. Odysseus consenting to his temporary imprisonment is one thing; but, as the tweaked example suggests, you cannot persistently frustrate someone's self-direction and demand their compliance by reference to the value of that same self-direction.

The third worry is related to this - and is, in fact, something that Wolff noted in his preface to the 1998 edition of *In Defense* (p. viii). There, he writes with some chagrin of having received congratulatory letters from various militia groups and survivalists after the publication of the first edition; these actual mountain men heartily endorsed what they saw as an obvious paean to individualism and rejection of public authority. It is odd that Wolff seems so puzzled as to how they could have gained this impression, for - some concluding remarks in the book about the value of solidarity and the need for a co-ordinating authority for our *political* autonomy (which concept he leaves fairly vague) aside - the conclusion that most obviously follows from the majority of *In Defense's* argumentation is that we have not only a valuable interest in, but a moral duty to maintain, an absolute right to reject any imperative which we have not endorsed through rational reflection.⁵ This simply rules out the possibility of anyone having legitimate authority over me without my explicit consent, and gives the mountain man all the theoretical ammunition that he needs to defend himself against the legitimacy of political obligation. If authority truly is the right to *command* and be obeyed, and the autonomous agent recognises no commands, then how can the two be reconciled?

largely what hypothetical contractarianism (e.g. Rawls 1971) aims to do - but it is important to note that there is no need for it to be; what matters is consent, and the features of the institution and the obligation are only important insofar as they affect whether we would or should (rationally, self-interestedly, etc.) agree to the political obligation 'contract'.

- 5 For the sake of clarity and fairness, it should be emphasised again that Wolff does not think we're forbidden from treating others as reliable moral guides, nor that others may not legitimately compel us to perform our moral duties (or compel us to refrain from breaching them). In both cases, though, the judgement of the others is a kind of shortcut or backup to our own moral sense - we may take "don't do that!" as a good reason not to do something because we think the speaker is morally reliable, but strictly speaking we act heteronomously if we refrain from acting just *because* they told us to rather than because we take their command as a good *prima facie* signifier of the action's wrongness.

1.2 Two Analyses of Political Authority: Conceptual Analysis

So far, the tale seems pretty bleak: authority as we have depicted it thus far just doesn't look like it can make demands of autonomous agents (at least, without compromising their autonomy). However, this is where the methodological turn becomes relevant.

In searching for a plausible account of political authority, there are two broad approaches we may take. One starts by taking for granted some pre-theoretic intuitions, tries to find necessary and sufficient conditions which capture those intuitions, and then turns to the political context to see how we might apply the concept. I call this the conceptual analysis or concept-prior approach. By contrast, the other - hereafter referred to as the functional analysis - starts with the political context, by considering what practical political role is played by a particular concept. It then generates necessary and sufficient conditions by reflecting on what has to be present for that role successfully to be played. The contrast between the two approaches can be captured by thinking about the difference between two questions: "Authority looks thus-and-so; how do we justify it in a political context?" and "What could we not achieve without a concept of political authority? How would the concept have to look in order to achieve these goals?". In this section, I provide a brief explanation of the problems of the first, concept-prior approach, by way of motivating a shift towards the second.

It is first necessary to explain why the approach holds its attraction. Generally, we have strong intuitions about what a political authority is, and what is or is not authoritative. For example, that authority must include an ability to issue commands rather than advice; and that these commands must give compelling reasons to action that are not simply prudential. "Do x or I will hit you" might give you a compelling reason to act, but we would want to avoid saying that it was an *authoritative* command. Rather, the reason has to be something like the avoidance of immoral behaviour - if I give you an authoritative command and you disobey, you have done something foolish, wrong, or both.

An account of authority which takes there to be a pre-existing concept with clear necessity and sufficiency conditions - the conceptual analysis approach - makes intelligible our feelings that “don’t kick that baby” is authoritative where “give me your lunch money” is not. When we distinguish between the two imperatives, the concept-prior approach explains, the second demand - for our lunch money - fails to meet the necessary and sufficient conditions (whatever they might in fact be) for being authoritative.

This approach can be illuminated with a parallel from epistemology.⁶ We might approach the concept of knowledge by taking as a starting point our intuitions that knowers ought to be right, and ought to be right reliably rather than simply by chance. Thus, something like the justified true belief account effectively gives us necessity and sufficiency conditions for a concept’s capturing our intuitions - “knowledge is this kind of thing [a correct belief about the world not arrived at by chance]; how do we go about justifying our claims to having it?”.

Notice that many of the historically noteworthy or dominant accounts of political authority are perfectly coherent when we take a concept-prior approach to authority. Hobbes clearly thinks that any rational agent has compelling reasons to accept *any* sufficiently powerful warlord. Rousseau famously argues that our rational self legitimates the operation of political authorities. Rawls tries to split the difference with a sort of “what any rational person *would* accept is permissible if actual agents mostly accept it” account, and so on.⁷

All of these accounts can be interpreted as adopting the view that authority is *x*, and the task is for us to work out how *x* can be suitably harmonised with other values *y*, *z*, etc.⁸ (By contrast, consider the functional analyst’s formulation:

6 See Craig (1999) for his thesis about what we would lack in an epistemological ‘state of nature’, where we had no concept of knowledge.

7 To return to epistemology, the parallels would be different accounts of what it is for a belief to be justified - they offer distinctive answers about what the necessity and sufficiency conditions are, but share the same presumption that anything which does in fact meet the correct conditions is what we mean by knowledge.

8 For example: Hobbes (1651) can be interpreted as making that case that authority is the right to command & be obeyed, and what gives an agent this right is a combination of power and rational self-interest on the part of others. Similarly, we can read Rousseau (1968) as using

“We have a practical project which makes *x* intelligible; is that project compatible with a separate project which makes *y* important to us?”).

And this, at first pass, is fairly compelling. We do seem to have some notion of authority to hand when we consider the command cases above, and it is from there fairly natural to conclude that the problem must be that the second demand fails some conceptual criteria for authority.

However, it is notoriously difficult to combine the pre-existing intuition of authority as “the right to command and, correlatively, the right to be obeyed” (Wolff, 1998) with a workable account of institutional obligation. For it seems that here there are only two ways something can be authoritative: either there are objectively compelling reasons already, or I have generated a reason (by, for example, consenting to the authority). But neither of these ways will work for political institutions, at least straightforwardly.

What these ‘objectively compelling’ reasons may look like is a difficult although perhaps not insurmountable problem. Raz (1986) and Darwall (2009) give two very different accounts of how we can be authoritatively bound by others even in virtue of reasons we may reject, for example, but in both cases overcoming the problem requires making the authority more or less straightforwardly moral. In other words, the authority of institutions comes ultimately from moral judgements; I ought or ought not to do something an institution tells me (not) to for precisely the same reasons as I would have if some morally correct passer-by gave me the imperative. This seems to leave the institution more or less out of the picture; its being formed for such-and-such a purpose, or having *these* particular features and lacking *those* others, is essentially irrelevant to its being authoritative. It seems fairly plausible, for example, that anybody who asserts we should not torture animals or infants is correct and in this sense authoritative - we ought to do what they say. But this directive could be issued by anyone or any institution without it becoming any less (or more) authoritative; whether it’s your granny, Holyrood or FIFA doing the directing makes absolutely no difference

the same concept of authority but justifying it differently - through the operation of the General Will and the rational self, and so on.

to our being obliged to obey. Clearly it cannot be the case that FIFA or your granny are political authorities, but if all that legitimises authority is the content of the directive, we have no way of individuating Holyrood as a (putative) political authority. So although providing objectively compelling reasons certainly have a role to play in an institution's being authoritative, it cannot be solely that.

Balancing political authority on explicit consent, on the other hand, leads us very quickly to problems of total voluntarism. If I should do x because I have consented to an institution governing me in the matter of x-ing, then it is not a straightforward moral claim - there is, by hypothesis, nothing independently wrong about (not) x-ing - but we famously do *not* consent to most of the claims and restrictions on us. Here, the institution in question gets the right to command (and be obeyed) only if I agree to be commanded; and this seems neither intuitively plausible nor practically workable.

To recap: the concept-prior account of political authority is appealing insofar as it matches with our intuitions about authority in other areas. In particular, it explains the parallels between political and moral obligations insofar as we tend to think that (for example) paying tax is, like giving a dehydrated person water, something we must do regardless of our feelings about it. It explains it, I suggest, by dint of eliding moral and political authority; transplanting moral authority, where the hard necessity and sufficiency conditions really do seem to simply be met by the correlative rights Wolff suggests, into an institutional framework.

If we take the concept of authority to come first, and require that political institutions meet the necessity and sufficiency conditions to claim authority, then we have an immovable conceptual object poised to resist any and all force we apply to it. Although the conceptual analysis approach allows us to interpret arguments about political authority, it does so in such a fashion as to limit the tools at our disposal - and, in doing so, makes the interpretation we get seem more inevitable than is in fact the case.

To give an example: presuming that political authority is a concept with these fixed conditions, which we then try and make fit with other values, will make it seem as if there is no way to reconcile authority and autonomy. If “the right to command, and be obeyed” is a fixed point, then it is difficult not to conclude that we have to sacrifice some measure of autonomy in order to accommodate the inescapable needs of authority; but, I suggest, the needs are only inescapable *because* we have taken authority to come first.

1.3. Two Analyses of Political Authority: Functional Analysis

Compared to the conceptual or concept-prior analysis, the functional analysis runs the state of nature experiments backwards. Rather than querying how we would justify such-and-such account of authority in some variant of the state of nature, the functional approach asks the question “what functions might we need a concept of authority for in the state of nature?”.⁹ The answer to this, and whatever the reasons we give for it looking this way, tells us a great deal about what it means for these institutions to be authoritative - and in turn, gives us criteria to assess the authority claims of particular institutions or systems.

It is worth noting in passing that functional analysis need not rule out any substantive proposal just because it is presented as a concept-prior account. Changing the second-order theory is not *necessarily* an endorsement or rejection of any first-order theory. Hauling out the epistemological analogy one final time, if we take the functional approach that we would lack the means to identify reliable testimony/testifiers absent a concept of knowledge, it might well be the case that justified, true beliefs are required to give reliable testimony. Similarly, perhaps a functional analysis of authority will reveal that Hobbes had it right all along. Importantly, though, he would be right in virtue of the values and ends we could not achieve without authority, rather than because he had given the correct necessity and sufficiency conditions for a pre-existing concept.

⁹ In addition to Craig’s “practical explication” approach to epistemology (op. cit.), an argument in favour of adopting the functional analysis in political philosophy can be found in Lane (1999).

An interlocutor could object that, at the very least, Hobbes and Rousseau *do* ask the question of what purpose authority serves: we have reason to obey the sovereign because we need to solve co-operation problems and brute force is the best way to do that, or we must obey the general will because so acting is constitutive of moral freedom; the shape of the authority is, in short, determined by what we want it to do.

This, however, is not quite right. The ways in which authority is justified are determined by which values the authors aim at, yes, but the concept of authority, *rather than its justification*, is taken for granted. That is, while Hobbes bangs the rational self-interest drum and Rousseau looks to moral liberty as giving us reasons to collectively self-legislate, they both presume the concept they are trying to justify - there must, they assume, be *something* out there which has a straightforward right to command (and demand obedience) from us, so the question is working out how we get there.¹⁰

Approaching the concept through functional analysis, with functional desiderata preceding and shaping rather than following from the account of authority, places vital questions front and centre.

An admittedly simplified example should illustrate the point. If we want authority to be simply “ability to command” - something like Weber’s (2004) notion of sovereignty as a successful claim to be the sole executor of legitimate violence - then there is no need for it to satisfy any requirements of justice, procedural consistency, and so forth. The effective warlord *is* authoritative according to this standard.

Conversely, if authority is supposed to carry some normative rather than merely prudential or instrumental weight (in other words, not just the ability but the right to command), then there will be extra conceptual requirements; the authority must be capable of more than brute coercion if the function it serves extends beyond the crude Weberian idea given previously.

¹⁰ See Hobbes (op. cit.) and Rousseau (op. cit.)

In the binary above, the obvious point on which to push the second account's proponent is *why* we want authority to have this normative capacity, especially since it makes it much harder to come up with a consistent and practicable concept.¹¹

There is a proliferation of ways in which we could answer the question, of course, but they are not important here. The point is that we encounter the question now, right at the start of our attempt to produce the concept. Instead of presuming that "right to command others" must be a coherent part of any notion of authority, and then running up against - for example - apparent problems of voluntarism, we drag the conflict into the open early on. This arguably complicates rather than simplifies matters, insofar as we cannot take a pre-existing mould of authority and see what fits neatly within it. It is surely, though, a better method than using the wrong mould; by inquiring why we want an account of authority, we have to be clear about what values will be served by such an account, and how.

The process thus articulates and distinguishes several important questions which might otherwise remain implicit within the discussion; further, it goes at least some way to providing answers to those questions.

If, for example, we want a concept of authority in order to serve the end of a patterned resource distribution, then there will be some potential features already ruled out. It seems unlikely, for example, that a concept which included or allowed for arbitrary changes on the part of the authority would be best for this role, since there are presumably fairly few patterns of distribution that are most effectively served by a whimsical and inconstant distributor.¹²

11 It is worth noting that at this point, conceptual analysis hits a brick wall: it, unlike the functional analysis account, cannot take the metaphorical step back and ponder which concept of political authority we should adopt (or which serves our desired function) to answer the question.

12 Interestingly, though, it's not immediately obvious that transparency (which is often taken to be a desirable quality) need play a major or even minor part in a concept of authority required for the "administration of things, rather than the ruling of people", although one could certainly make a case for it. If we want authority to serve a more 'positive' role, in the sense of enabling or creating goods (rather than simply shuffling them about or making them

There are two values, or types of value, that seem best served by functional analysis. On the one hand, we have the methodological advantages: as applied here, functional analysis is explicitly a tool to reconcile autonomy and authority. As such, it is purpose-built to defeat the problem of the mountain man, whereas the concept-prior approach must simply shrug and tell us to embrace or ignore him.¹³ Relatedly, functional analysis does not make unattractive theories or dilemmas seem inescapable - the reason that concept-prior accounts of authority give us a binary choice with regards to the mountain man is because of their own presuppositions rather than because they have captured some unavoidable truth about the world.

But this alone may not be enough to motivate functional rather than conceptual analysis. Apart from the aforementioned fact that functional analysis often seems to result in far more complex explanations, which on the face of it must count against the theory, the value of practical explication is that it most closely approximates something like materialism (in the Marxist sense) about political philosophy. The proper aim of political philosophy is, surely, to think seriously about what humans (and other social agents) need and ought to have, and - inconveniently - these needs as well as the methods by which they can be fulfilled tend to change temporally and geographically, politically and economically.

We should therefore consider what purpose authority is supposed to serve *before* deciding on its features, rather than importing pre-theoretic intuitions and trying to justify them; and we should be aware that our being able to make

easier to get at), such as being a facilitator and marker of community decisions, then a stronger transparency requirement looks probable.

¹³ To clarify, I don't take the functional analysis to be a way of mollifying the mountain man. It's not supposed to convince him that's he's wrong about what we have to believe about authority if we value autonomy (though it might do so), but to show that his presuppositions aren't inescapable: we can, *contra* the mountain man's claims, perfectly coherently value autonomy and authority, so long as we're clear about what roles they're serving. If the mountain man accuses me of shifting the goalposts (as Cowan has suggested he might), it will be a charge I'm happy to accept - since part of the *point* of the functional analysis is to show that we can and perhaps should shift the goalposts, depending on the kind of game we want to play.

sense of particular accounts of authority via conceptual analysis does not in fact mean that authority is an uncontroversial descriptive concept.¹⁴

To return to the intuitive characterisation of authority as providing compelling reasons to act given at the start of this section: the functional analysis need not, and probably won't, reject this desideratum for political authority. What it does suggest is that the nature of those reasons is far more dependent on substantive arguments about value - procedural consistency versus material impact, co-operation versus competition, and so on - than we might anticipate. In short, the purpose of the functional analysis is to put the values or goals we want authority to serve first, and to construct the concept with those values in mind, rather than take the concept to be set and then try to make its justification line up with our values. If it turns out that, in fact, there really is no way to make a plausible concept of authority which can serve those values (whatever they are) then the mountain man will be vindicated. For now, I merely offer the methodological and practical reasons to adopt a functional analysis, returning to the question of what role we actually want authority to play in Chapter 5.

2. Autonomy: Criticisms and another Functional Analysis

One possible response to the problem of the mountain man, and particularly to the problems of justifying democracy through autonomy, is to accept that autonomy is not the only value. Where autonomy and authority clash, autonomy may come off worse, but in the process equality (for example) may be increased. The pluralist about value can thus adopt a fairly modest strategy of concession. I do not address this strategy for two reasons. First, and for reasons that will be laid out in more detail over the course of the thesis, it does not seem to me to involve the correct conception of autonomy. Second, it is not a defence against the stronger - and, I think, more interesting - version of the autonomy criticism.

¹⁴ Here I am to some extent standing alongside Finlayson's criticism of the notions of reasonableness and neutrality in political philosophy, where she argues that "politically neutral 'common sense'" usually turns out to be both politically controversial and inherently reactionary. See Finlayson (2015, p55).

That is, the weaker form of the criticism says merely that autonomy is not the only value, but the stronger variants argue that it is actively *disvaluable* - we should not be concerned with clashes between autonomy and authority because we should not be concerned with autonomy in the first place.

I leave aside the group of criticisms focussed on the purported non-existence or inaccessibility of the self, and also ignore the objection that self-determination is impossible because causal determinism is true (that is, we cannot be autonomous if this is understood as having some measure of control over our actions, since our actions are entirely dependent upon long causal chains stretching back to the Big Bang). Instead, I consider two arguments to the effect that deploying the concept of autonomy is harmful in theory or in practice.¹⁵

2.1 Autonomy is Conceptually Unappealing

The first criticism, that autonomy is actively disvaluable at a conceptual level, has been levelled by a number of critics, both communitarian and feminist. I address the feminist charges only, as the communitarian critique is more closely bound up with ideas of self that lie outwith this chapter's purview. The meat of this charge is summarised by the worry that the ideal of autonomy

nurtures a stark individualism fueled by the silent assumption that autonomous man is free to sidestep the constraints of materiality and the power of social-political structures in his projects of radical self-making.¹⁶

The pursuit of autonomy is thus cast as an attempt to cut oneself off from personal relations and responsibilities; to be autonomous is indeed to be the autarchic mountain man - and this is precisely why we should abandon the notion. We can cash this thought out in two ways. The first is that, as an androcentric ideal, autonomy should not and cannot be esteemed as a purportedly universal value. The related second conclusion is that pursuing such an ideal is anti-feminist (and/or, given the traditionally Western and Anglophone emphasis on this area of philosophy, colonialist or imperialist) and thus actively

¹⁵ Both these sets of criticisms (along with the critiques that I *do* consider) are given a more detailed airing and then rejected in, among others, Ch. 2 of Friedman (2003).

¹⁶ Code (2000, p183).

harmful to any oppressed or victimised group. To draw a parallel from elsewhere in political philosophy, we might object to the notion of negative liberty by arguing that freedom *from* is, or has been, primarily the concern of the privileged and/or wealthy precisely because the primary concern of capital accumulators and dominators is removing limitations on their exercise of power. To presume the importance of unfettered choosing over safety, health and so on is simply to apply an inappropriate standard, and one that - deliberately or not - excludes those whose main concerns are not the absence of external legal restraints, but the presence or otherwise of basic social and economic provisions.

It is unarguably the case that a significant proportion of theorists concerned with self-direction have expounded precisely the kind of hyper-individualistic account that Code charges them with.¹⁷ When Wolff says that the “primary obligation of *man* [italics mine] is autonomy, the refusal to be ruled” (op. cit., p18), he very obviously has in mind the sort of self-sufficient rational chooser that stalks the halls of political philosophy. The idiosyncratic notion of autonomy at play in (for example) reducing welfare spending in order to make folk self-sufficient equally clearly draws on the caricature of the autonomous agent as unhindered by ties of dependence or mutual support.¹⁸ Further, given the close historical association between gender roles and particular values - the view (most commonly found in the columns of the right-wing *Daily Mail*) for example, that autonomy-as-self-sufficiency is a valuable trait for male-identified agents but positively poisonous for female-identified agents - then, if this is indeed what is meant by autonomy, we should join its critics in repudiating it as a useful normative concept.

I argue, following Friedman et al, that the *concept* of autonomy need not carry any such baggage, and that it would be a mistake to conflate the concept with some of its particularly noxious conceptions. All that we have said of autonomy thus far is that it requires power and authority over oneself.¹⁹ These features are

¹⁷ See, for example, Frankfurt (1971), Nozick (1974), Rand (1982) and Stirner (op. cit.). As suggested above, Wolff (op. cit.) probably falls into this category too, albeit reluctantly.

¹⁸ That, at least, is the most charitable interpretation one can give.

¹⁹ Not political authority over oneself, however that would work, but in the sense that we are competent reasoners, and our wanting to do something is taken as a strong though *prima facie* reason for us to do it. I discuss this in much more detail later, but for more on autonomy

in no way necessarily linked to ideals of being an unfettered chooser. To give an example, someone with a family could be autonomous despite (or, as I argue later, because of) the various responsibilities placed upon them by these relations, and without at any point having consciously decided to undertake these responsibilities. Their power over themselves, without getting *too* deeply into competing conceptions at this early stage, does not require that they be free to pick any option at any time. Similarly, they meet the authority condition so long as their status as an agent is taken to demand recognition, rather than only if it's the case that their word is (metaphorically) law and cannot be gainsaid.

Still, the autonomy sceptic might yet be uneasy. However one dresses up these ideas, they are still, overtly, concerned with exercising (perhaps imposing) our wills. Is it not merely pushing the question back a stage to assert that whatever autonomy consists in, it must give us power? By analogy: say that I am trying to convince someone to buy a car, because of the increased ease of movement this gives them. If they are apathetic or hostile to this, perhaps (to try and keep the parallels fairly close) because the purported "joy of the open road" is a paradigm of individualistic self-sufficiency, then it serves no purpose for me to point out the multifarious different ways in which increased freedom of movement can be applied. If the open road's joy is no object of value, then offering up ways in which it can be accessed by a wider group or engaged in new ways is just re-arranging deckchairs on the Titanic.

One strategy that would be open to me, however, is to argue that the joy of the open road is not the real value at all - it may be widely taken to be so, and this widely-accepted notion could well be oppressive, but the actual value of car ownership is something entirely different. Perhaps the value lies in the fact that I can now meet with friends more easily and often, or that new job opportunities have opened up, or - bleakly - that I have some avenue of escape.

How can this strategy be applied to power and authority? Doing so fully requires laying out a conception of autonomy, but some progress can be made at the bare

concept stage. Clearly, power over one's life *can* be interpreted in a hyper-individualistic way: as being the sole director of one's actions with no regard for responsibilities or relations. Cast in such a light, autonomy is simply a fantasy, if an unhealthy one - no-one can avoid the "constraints of materiality" or the inconvenient fact of social norms. But *need* it be interpreted this way? Surely not: demanding the power to self-direct is simply to demand freedom from domination, and this is something that coheres well with feminist attitudes towards employment and reproductive autonomy (to give but two examples). Dependence on another, to take the opposite tack, is not necessarily a reduction of power to self-direct *unless* that dependence has some oppressive aspect. Being dependent on a friend for emotional support, or on a carer for some physical tasks, does not render an agent non-autonomous unless, say, that friend uses this relation to malignly interfere with the way my life goes (imagine a case where they use emotional blackmail to prevent me from a taking a job or maintaining a relationship with someone else).

Similarly, holding that we must be authoritative over our own lives is not equivalent to holding that no one may challenge our behaviour, or that we cannot be beholden to others. If, for example, I have decided that I want to go to the pub tonight, but someone reminds me that I already promised to dogsit for a friend, then my being under obligation doesn't undercut the authority of my previous plan: the objection is not "you aren't an appropriate source of values regarding your life", but "you have already decided that dogsitting is in some way a valuable activity (and made a commitment based on that which it would be immoral to abandon)". While my freedom of action is plausibly being restricted, the restriction is vastly different from that put in place by oppression.

By way of example: in the UK, heterosexual cisgendered men face no establishment pressure or questioning over their sexual preferences or identity (beyond perhaps the lingering Victoriana of 'settling down', marrying and starting a family being regarded as the default). They are held to be entitled to their attractions, and it would be widely regarded as absurd to inquire as to why they feel this way. There are of course occasions where their behaviour can be challenged, and being authoritative over their attractions manifestly does not

equate to a right to have them required, but, crucially, *that* they have these attractions is taken as *prima facie* reason enough for them to be treated with (some fairly minimal level of) respect. They are, in this sense, authoritative over their sexuality.

Conversely, and to take a recent case, the homophobic and transphobic Section 28 which used to operate in British schools stood as a direct challenge to the authority of LGBTQ agents over their sexual identities. The (barely) implicit assumption about agents being made by such legislation must be something like: *that* they have these attractions or this identity is reason enough to prevent them, *with the coercive force of the state*, from being acknowledged in education as living a life as valuable and ‘normal’ (or, indeed, abnormal) as any other’s. This was not primarily restriction of action - it certainly indirectly results in restrictions of action through maintaining the heteronormative structure it was designed to defend, that said - but an obvious refusal to grant the agent status as a source of values; their having particular desires or identities was not accepted as reason to treat these desires and identities with respect, despite there being no normatively important difference.²⁰

While dogsitting can represent a restriction of action and yet have no effect on my authoritative status, thereby showing that authority does not require unhindered and unchallengeable choosing, the Section 28 example highlights that it is not possible for us to be authoritative over ourselves without our desires and identities being presumed to have normative force. To put it another way, the authority in question does not require that we may brush off any and all demands for justification; it *does* require that those demands be addressed to us, as agents and independent sources of value, in the first place.

I believe this goes some way towards dealing with the criticism that the concept of autonomy is unduly individualistic and it is therefore a mistake to pursue it. Indeed, I would argue that something very like the power and authority conditions are implicit in feminist accounts of what is necessary for living free of

²⁰ This does not, of course, imply that there are no differences in the way that people of different identities experience their lives - simply that there is no defensible normative basis for acceptance of one set of consensual relations and hostility towards another where the only distinction is the sex/gender of the partner(s).

oppression.²¹ The other criticism remains, however - the objection that while the concept might not be harmful, its pursuit is in practice either a distraction from more pressing matters, or apt to worsen the situation of those who pursue it (and/or people close to those who pursue it).

2.2 Autonomy is Disvaluable in Practice

This criticism harks back to the point made earlier about the wielding of self-sufficiency, “pulling oneself up by the bootstraps” and the like in the service of public spending cuts: whether or not autonomy *requires* that we all do the dance of the mountain man, it seems at least to allow it. Given this, and given the pre-existing political and social structures which tend to conflate autonomy with autarchy and/or self-mastery, the critic might object that however conceptually innocent it is, the practical upshot of asserting autonomy as a primary value will be malignant.

Return to the car example mentioned previously. I could make a convincing argument to the effect that there were various genuine values to possessing a car - that the concept of having a car is not, in fact, a necessarily harmful one, and even a worthwhile one - and still be blocked off by objections that for various socio-economic reasons (pollution, the relative cost of owning and running a car, etc.), car ownership is not a practically worthwhile goal. Or, to use a political example which is closer to the bone, we might think that there are good philosophical reasons to adopt unanimous direct democracy but pursuing it would be impractical, or there are more pressing issues to organise around, and so on.

The objection has three strands:

1. Autonomy is conceptually *and* practically valuable, but given prevailing social and material conditions, there are more important tasks to be addressed.
2. Autonomy may be conceptually valuable, but its pursuit under current social and material conditions is likely to harm the autonomy seeker.

²¹ See Friedman (op. cit.), Oshana (op. cit.) and Stoljar (2000).

3. As 2, but its pursuit is likely to harm those close to the autonomy seeker.²²

Taking the strands in order, 1 can be partly answered with the stereotypically philosophical “yes, but on the other hand...”. To argue that, for example, personal-political autonomy must take precedence over survival, or survival at above a bare subsistence level, would be to ignore the basis of autonomy - we cannot respect agents as sources of value and simultaneously hold that they may not prioritise their survival over theoretical commitments.²³ But on the other hand, whether autonomy is a value prior to (say) personal security, or even a value that includes it, will depend on the particular conception of autonomy that we use. Similarly, a suitably social account of autonomy will place great emphasis on addressing conditions of oppression and exploitation rather than advancing the individual through or around such barriers - that is, the pursuit of autonomy (properly conceived) will be, at least in many if not most cases, concerned with socio-economic liberation before it is concerned with maximising the freedom of action of one person, or of arbitrating between the competing ends of already-fairly-autonomous agents. The response, then, is a dual one of conceding that autonomy may not be the sole or the primary value, but also that a suitable account of autonomy might very well turn out to include many of its purported competitor values.

The second part of the objection, that pursuing autonomy may be harmful to the would-be autonomous agent, admits of a parallel response regarding the distinction between the standard or popular concept of autonomy (as something akin to self-sufficiency or self-mastery) and the reconceptualised idea of autonomy as power and authority. Just as the notion of empowerment is multiply traduced and used to justify everything from privatising medical care to cutting disability benefit, but remains a genuinely worthwhile principled and pragmatic goal if properly constructed, so autonomy is only likely to be harmful

22 The important distinction between “autonomy may be bad because harmful to the seeker” and “autonomy may be bad because harmful to those around the seeker” is drawn in, e.g., Friedman (op. cit., pp41-43).

23 Although this is to some extent dependent on what precisely those commitments are - perhaps I may prioritise survival over a commitment freedom of expression, but not over a commitment to treat others as ends in themselves.

if mistakenly conceived. If someone pursues self-sufficiency, with the attendant cutting off of important social relations, or self-mastery, with the accompanying denial of social and material conditions' effect on identity, then it *does* seem likely that they have been harmed. Nevertheless, as has been argued before and will be argued again in this chapter, self-sufficiency and self-mastery are *not* autonomy. To put the response in the simplest terms: the potential harm of pursuing a misconceived idea of autonomy does not mean that autonomy is itself potentially harmful.

Finally, there is the claim that pursuing autonomy may have deleterious effects on those close to the autonomy seeker.²⁴ The objection can in some cases be rebutted by a version of the same response as above, but this rejoinder will not work universally. While autonomy *need not* invariably involve the abandoning of pre-existing relations, there seem to be no principled reasons to suggest that it *cannot*, and in some circumstances, such as those where the relations are exploitative, then we might well return to the view that autonomy endorses removing oneself from them. Take the case of someone who decides that their self-conception is inconsistent with living in a densely-populated city, and what they really want to do is to go and live in the woods somewhere. They need not be swinging for *Walden* (particularly if they have no aunts to bring them food every weekend); they can quite coherently value the friends and experiences they have accumulated in the city, but simply feel that, for whatever reason, they would feel “more themselves” and/or more in control of their lives in a rural area.

So far, their attitude seems perfectly consistent with autonomy conceived as power and authority, but it is clearly plausible - indeed, likely - that others will be negatively affected by this decision. Friends will miss them, flatmates will need to engage in various tedious administrative tasks; depending on their job,

²⁴ I have chosen not to focus on starker cases of self-direction leading to the harming of others on the basis that such actions are independently ruled out by moral concerns - no-one, save perhaps Stirner, would defend the idea that autonomy can countenance a right to unprovoked murder, given that (even *if* such acts were internally consistent with maximal equal autonomy, something I dispute) this would very obviously be immoral for more reasons than that it ignores the autonomy of the victim. Less parodical cases of this type, it seems to me, do a better job both of bringing out the critic's legitimate worries about fetishising the individual, and defending autonomy against such worries.

the move may even have a damaging impact on those they work with (perhaps they are a doctor who subverts or refuses to comply with unjust disability assessment schemes, and their replacement will be a running-dog motivated by the profitable prospect of declaring as many people fit for work as possible). It appears, then, that at least the charge that autonomy causes or may cause harms to others will stick - and if we are engaged in defending a concept which benefits the individual but harms others, are we not veering dangerously close to precisely the kind of right-libertarian land of take-what-you-want that I have been at pains to reject thus far?²⁵

There is certainly something intuitively unappealing about the thought that supporting one person's self-direction may be to deny another some non-trivial benefit - but this is an intuition of a kind with hesitating to publicly condemn an incompetent doctor because they seem pleasant. It may not be *nice* to do so, but then - as TV Tropes would have it - good is not nice, and people blamelessly hurt each other all the time. If one person in a relationship comes to realise that they are no longer happy to be with the other(s), then revealing this information, and ending the relationship, is highly likely to be incredibly hurtful no matter *how* gentle the let-down is. We cannot, however, seriously entertain the idea that the relationship-ender has done something wrong here - as Friedman says, "a person has no general responsibility simply to remain in just any social relationships in which she finds herself."²⁶ We obviously bear some basic responsibilities towards each other which limit (to some extent) how and when we may break off certain relationships; I cannot abruptly decide that I don't want to see my partner any more if we're alone in the wilderness and I have the only map, for example. Here, though, I am under general moral obligations - the problem is not that my partner will or may experience distress at the news I don't want to be with her, but that I will have essentially left her to die of exposure. Again, the issue is not specifically that my pursuit of autonomy has harmed another, but that I have broken a (particularly strong) moral obligation.

25 "Eve: All this riot and uproar, V... is this Anarchy? Is this the land of do-as-you-please?
V: No. This is only the land of take-what-you-want."

See Moore/Lloyd (1982).

26 Friedman, op. cit., p42. See also Humphries (forthcoming).

It is not my being powerful and authoritative which has caused harm to another: it is my using this power inappropriately.²⁷

But failing to maximise somebody's happiness is not the same thing as failing an obligation (unless one is an act utilitarian, of course). I am not failing my obligations or duties of care to a competent adult by deciding that I need a better social life, or to return to university or the like, even if we very unsafely assume that the other person will have a marginally less satisfactory life as a result. Similarly, if I know that a friend would be pleased to see me today, but it is vital that I get some research done - or even if I have a strong desire to take the day off to eat soup - then presuming that I have not already promised them I would visit them, going to work (or eating soup) doesn't constitute my wronging that friend. To respect another as a source of values is not to regard those values as always taking priority over one's own; it is simply to acknowledge that they *are* values, and to give them appropriate consideration. Absent some previous obligation to the contrary, we *can* blamelessly make people unhappy, and the possibility that the autonomy seeker will harm others around them is one that must be accepted. To argue the contrary is, essentially, to argue that we cannot change the relations we stand in without the explicit permission of all other parties, a deeply reactionary stance. It is also a stance that is plainly inconsistent with the value of self-direction understood as being powerful and authoritative over our lives.

It is worth making one final point on this matter. To say that we are blameless "absent obligation" is not to suggest that only explicit, contractarian-type agreements (between rational individuals, presumably) can put us under obligation. Equally, the question of whose value should take priority is difficult to give an abstract, principled answer to, but there are paradigm cases showing that it is certainly possible (and sometimes fairly easy) for us to weigh up these values. If, in the example above, I know that my friend is lonely and depressed, then staying at home to eat soup will be blameworthy irrespective of any claims that "I never *said* I'd go and see him", and of the fact that I genuinely enjoy soup. Recognising them as a source of values *does* involve recognising that (on

²⁷ With great power comes great etc.

this occasion) their wishes are more important than my own, and insofar as our relationship constitutes an endorsed part of our identities, this obligation has no malign effect on our autonomy - indeed, it is actively beneficial to it.²⁸ As such, it misses the point to say that autonomy is an overly-individualistic or oppressive concept because it *may* sanction courses of action, such as the ending of particular relations, which bring harm to others. For one thing, autonomy (as I conceive it) is perfectly happy for a deeply socially-embedded agent to be autonomous not despite, but *because of* their network of obligations; for another, if we are not prepared to sanction the occasional harm to others, we are plunging head-first into exactly the kind of dominating, harmful-norm perpetuating ideology that autonomy's feminist critics would reject.²⁹

2.3 Recap and Another Functional Analysis

This subsection has addressed the objections that the concept of autonomy may be inherently harmful, and that its pursuit may be harmful in effect. The first objection, I argue, mischaracterises autonomy as being necessarily the ideal of the hyper-individualised, self-mastering rational agent who is beholden to nothing and no-one. While this is unarguably a fairly accurate picture of a number of conceptions of autonomy, it is not a required feature of the concept itself. All that is conceptually necessary for autonomy is power over one's life and authority (weakly construed for the moment as an agent who is self-respecting and regards their self-direction as generally reflecting their values) over it, and neither of these conditions is inextricable from oppressive conceptions of autonomy.

The second objection amounts to three slightly different criticisms: first, that autonomy is valuable but only one value among others; second, that pursuing autonomy is liable to harm the pursuer; and finally, that someone seeking to increase their autonomy is liable to harm others around them.

28 Of course, whether or not they are depressed, then I may equally go and see them ('sacrificing' my research or soup in the process) without any autonomy loss - again, the relationship makes up a valuable part of my self-conception. The point is simply to show that moral obligations can play important roles in autonomous lives.

29 For a parallel and relevant criticism, see, e.g., Bartky's (1990) objection that an ethic of care looks dangerously like a re-affirmation of traditional 'good woman' roles. See also Kaplan (1994).

In response to the first criticism, I suggest that while autonomy may turn out to be a far wider-ranging value than previously thought, there is nothing in the ‘thinner’ version to suggest that we must at all times prioritise autonomy over survival, or other basic goods. The second prong admits of a mischaracterisation defence similar to the one given in the previous paragraph: it seems entirely plausible that people could bring suffering and/or harm on themselves through pursuit of *purported* ideals of autonomy, like self-sufficiency or disassociation from involuntary relations, but all this tells us is that the concept of autonomy has been bastardised into something like autarchy or free choosing - not that autonomy, correctly conceived, need be harmful to its seeker.

Finally, the thought that the pursuit of autonomy may be harmful to others needs a more complex response. Autonomy does not licence immoral behaviour. Even if it were the case that my autonomy would be increased or improved by setting fire to a noisy neighbour’s flat, I may not do so on pain of (apart from anything else) failing to accord them the same status as a source of values that I claim for myself. Autonomy may still licence behaviour that causes others upset, distress, or even harm. But the alternative suggestion, that we must remain in any relationships which would distress others if ended, is clearly an indefensible position and one distinctly at odds with the historical and political-philosophical commitments of feminism. Just as an autonomous agent may nevertheless be deeply unhappy, making someone else unhappy is not, in itself, a violation of their autonomy.

I have, hopefully, defended against claims that autonomy is necessarily disvaluable or harmful. Before concluding the chapter it is worth drawing out the positive case for valuing it, through a functional argument which has been immanent in responding to the foregoing criticisms.

The case is this. The concept of autonomy, of being powerful and authoritative over our lives, is valuable because it allows us to describe and explain the importance of self-direction. We do not, after all, desire to simply be ‘wanton’ choosers settling on options more or less at random. The idea of autonomy as being in some sense authoritative over the decisions that we make illuminates the notion that self-direction is valuable because it reflects (or instantiates, or

whatever we wish to use at this early stage) something valuable about human agency, and agency requires more than random selection. This, of course, invites the question of why agency is valuable at all. One option would be to claim that its value is a primitive, akin to the way that utilitarians sometimes talk about the value of pleasure. I'm somewhat sympathetic to this, but my view is closer to Marx's notion (2010, pp86-92) of "species-being" (although, to be clear, I do not adopt the idea wholesale). This is in part the idea that humans' unique or essential property is that we consciously and creatively shape the world around us in accordance with our values and desires. Agency is thus valuable because and when it allows us to engage in this creative activity; as a useful corollary, this also explains why autonomy as being powerful and authoritative captures the value of agency where random selection does not - mere 'plumping' does not evince the idea of conscious creation that Marx appeals to.³⁰

Similarly, autonomy must involve some exercise of power if it is to be able to explain what is gained by successful self-direction, and lost through such direction being frustrated. If we imagine the second, kidnapped Odysseus, it is his loss of power which accounts for our feeling that he is not self-directed: whatever he has reflected upon, decided upon, and so forth, that decision is not mirrored or even acknowledged in his actions (or lack thereof). Autonomy need not, as I've argued above, be an explanatory factor for the value of *unrestricted* power or complete self-sufficiency - apart from anything else we have autarchy to hand if we want to make those kinds of arguments - but it doesn't seem consistent with autonomy's role that an agent could be completely unable to self-direct and yet be autonomous.

The functional analysis of autonomy, then, suggests that power and authority are required features of a concept which seeks to explain or make intelligible the practice and value of self-direction. It is *not* required by this concept that we be unrestricted choosers nor rugged individualists - in fact, as I argue in Chapters 3 and 4, the most plausible conception of autonomy is one which takes very seriously our social embeddedness. In the next section, I perform the necessary groundwork to argue for such a conception by showing that the concept of

³⁰ Someone might then inquire *why* creatively working over our material conditions is valuable. This, I suspect, is the point at which I give up and claim that the value is a primitive.

autonomy which plays the functional role required will not simply be given to us by adopting moral autonomy.

3. Acting Responsibly and Acting Autonomously

Given the importance I have placed on the concept of autonomy as something which helps to explain the value of “living one’s own life” or acting in accordance with deeply-held principles, it is worth taking some time to head off the reasonable thought that I might simply be talking about moral autonomy: that the concept which best explains why we want autonomous agents to be powerful and authoritative is the same concept which explains the importance of making moral decisions in the right way. In other words, there is a potential misunderstanding here that I wish to avoid, namely that valuing autonomy because we value agency suggests that I should give an account of *moral* autonomy. This section is devoted to avoiding that misunderstanding by showing that although the concept of moral autonomy does call on the value of agency, it does so in a way which is importantly different from the way that personal (or, foreshadowing Chapter 3, personal-political) autonomy does.

I argue that moral and personal-political autonomy are distinct concepts playing different roles in our philosophical discourse, and that moral autonomy requires (morally) responsible reasoning where personal-political autonomy has no such requirement. Finally, I provide several arguments to the effect that an agent can be morally responsible without being powerful and authoritative over their lives (in the ways necessary to explain the value of self-direction), and vice-versa.

To begin with a straightforward pair of definitional claims:

Moral autonomy is concerned with what we owe to others, and only weakly sensitive to our desires. We are compelled on pain of conceptual confusion towards adopting certain attitudes (taking certain actions, living a certain way) just by dint of the capacities we must use to consider them.

Personal-political autonomy is concerned with what others owe to us, and is strongly (although not indefeasibly) sensitive to our desires. Certain of our normative attitudes may be inconsistent with our own desires, but this is owing to the practical contradictions implied by the two sets of beliefs rather than any conceptual incoherence.³¹

Since moral and personal-political autonomy aim at different goods, we have different failure conditions for them. It is unlikely that moral autonomy can perform the functional role specified in the previous section. To support the stronger claim that the two are conceptually independent, I must show not only that there are conceptual differences and different failure conditions - for as I have conceded, there are also conceptual similarities and shared failure conditions - but that we can have one entirely without the other.

One noteworthy feature shared by all non-Kantian accounts of autonomy is that two agents can make morally autonomous decisions regarding the same question and come up with different, even mutually exclusive answers without either being wrong or heteronomous: autonomy is not necessarily a matter of coming up with the right answer *de re*, but of taking the right approach to the question. While there might be cases where restrictions of rationality, forward planning and so on mean that *in practice* there is only one right answer, these will be exceptions rather than the rule. Thus, Annie autonomously decides that x-ing is wrong, while Brendan equally autonomously settles on the position that x-ing is permissible; this may be due to differing higher-order ethical commitments, views of moral psychology, or simple probabilistic judgements (to give three out of a much wider set of possible commitments) and, so long as none of these other commitments themselves fail whatever restrictions are taken to be in play,

³¹ I argue later that it is possible for an agent to be conceptually confused about whether such-and-such a course of life can be defended as autonomous, but as suggested on the previous page, we can assume for now that there is no reason internal to such a conception of autonomy that anyone should *need* to defend their self-direction in terms of autonomy - someone who self-consciously commits to an elitist or egoist view can (and probably will) fail to be autonomous, on my view, but this is because of what autonomy actually requires, rather than because they are committed to courses of action which are inconsistent with some idea of autonomy that we have stipulated they are somehow also committed to. See Chapters 3 and 4.

Annie and Brendan's disagreement suggests no failures of moral autonomy on either of their parts.

Further, we can judge not just idiosyncratically, but *wrongly*, and still be autonomous. Imagine that Brendan's probabilistic assessment is simply incorrect, and correcting this assessment would cause his view of x-ing to align with Annie's. So long as, again, his mistake is not the result of slapdash (arbitrary/culpably ignorant/etc.) reasoning, Brendan's decision is still autonomous; and this does seem to accord with our intuition that one cannot require near-omniscience in order to be autonomous.³²

It also, however, begins to make moral autonomy accord closely with another normatively weighty notion, that of responsible moral action. In fact, it seems that the morally autonomous agent must also be the morally responsible agent; that, in Oshana's words, "autonomous agency seems to be both sufficient and necessary for responsible agency" (2002, p262).³³ It is to this claim that I now turn.

In the next subsection, I argue that being a morally responsible reasoner is a necessity condition for being morally autonomous, and in the following subsection provide several cases where meeting this requirement of moral autonomy (that is, responsible reasoning) is independent of meeting those of personal-political autonomy.

3.1 Moral Autonomy and Responsible Reasoning

The first thing to do here is to be clear about what we mean by moral responsibility. An obvious interpretation is something like "causal responsibility for a morally-loaded action" or "appropriate blameworthiness for some action" - I am morally responsible for causing a cat suffering because it was I who deliberately kicked it and thereby created its suffering, and so on. This is not,

³² See, for example, Levy (2009).

³³ To be clear, Oshana denies that these necessity and sufficiency conditions obtain: but she denies the necessary linkage between *personal-political* autonomy and responsibility, rather than moral autonomy and responsibility (and indeed I make the same denial).

however, the sense in which it is used here: rather, what I mean by responsibility is the act (or state, etc.) of acting or judging in a responsible fashion. Responsibility (for the purposes of this discussion) should thus be thought of as denoting a particular quality or virtue of an agent instead of suggesting a particular causal link - we are after the agent who is responsible in the way that a pipe-smoking boat captain is supposed to be responsible, rather than the way that the accused is supposed to be responsible.³⁴ As such, I will refer to responsible action or responsible reasoning, rather than to moral responsibility.

We can thus largely leave to one side tricky questions about folks' metaphysical ability to do other than they did. The crucial conceptual features here are not (or not solely) whether it was possible for me to act differently, but whether the act I *did* perform (judgement I made, etc.) meets certain internal and external criteria regarding the thought process that led to it. To act responsibly is to take certain relevant reasons as motivating and to act according to some appropriate set of principles (and to be willing and able to reflect critically on these principles), just as the concept of moral autonomy is concerned with the formation and adoption of specific (albeit specified at a second or higher order) principles or reasons for action. Below, I suggest some non-exclusive contenders for what may be required to set and meet such principles for responsible action; that these requirements will apply equally well to extensions of the concept of moral autonomy elucidated in the previous section is not coincidental.

Speaking very broadly, the first requirement of morally responsible action is a particular form of self-consciousness. In order to act responsibly, we should be aware that our (apparent) agency imposes requirements on us, and also aware of the rough limits of these requirements. For example, I act responsibly if I realise that my being able to seriously injure or kill someone with a car means I must

³⁴ Again there is likely to be some overlap - if someone is literally incapable of being responsible in the former sense, we would be unlikely to suggest that they are appropriate subjects of blame or reprimand (although they could obviously still be strictly causally responsible for something). Conversely, failure to reason responsibly before action, absent excuse, is one good reason to hold the agent blameworthy for the moral disvalue of the result. Further and more detailed discussion of responsible action, and of moral responsibility as holding responsible can be found in, among others, Fischer & Ravizza, (1991, 1998); Levy (2009), Oshana (1997), and Steadman (2012).

think carefully about where and how I drive, but it would be at least *prima facie* strange for me to worry about whether the state of pedestrian crossings in Bristol means I should not drive to Edinburgh, and absent other factors no-one would be inclined to think that my worrying so indicated a particularly virtuous level of responsible reasoning.³⁵

Of course, in just the same way that one can act or reason wrongly but autonomously, we can act or reason wrongly but responsibly. Taking the driving case above, imagine that I have no good (internal) reasons to thoroughly check my car's brake lines, with the result that halfway down the M8 I cause a pile-up through being unable to stop in time. If we stipulate that the responsible driver can *never* be the cause of an accident, then I will be judged irresponsible, but again very few people would argue that we ought to exhaustively check our cars before setting off in order to be responsible. It seems, again, as if the fact that I performed all the checks I took myself to be obliged to perform, and was non-culpably ignorant that there were other checks I should have performed, means I have acted in a morally responsible manner (and, curiously, might mean that I am *not* morally responsible in the other, apt-for-blame, sense).

Further, and again with near-perfect isomorphism to thinking autonomously, we can think responsibly about the car safety question and come up with different answers: Anne argues that all things considered, there will be more harm than good as a result of mandatory rigorous pre-drive testing, while Brendan makes the case that tougher restrictions on private motor vehicles might be no bad thing. Neither of their claims need be made in bad faith, or on arbitrary grounds, and so on, and so we have the third parallel between autonomy and responsibility: the possibility of blameless disagreement.³⁶

35 If I just considered the possibility and dismissed it (rather than being swayed by it, or it not coming to mind at all), an observer might perhaps be happier to say that I was being exceeding responsible, but even then the verdict seems fairly arguable.

36 It also seems possible at first glance that we could reason both responsibly and immorally (Anne earnestly believes that increased traffic accidents will toughen up humanity, for example); whether immoral action requires some failure of responsible reasoning is, again, an interesting tangent but not one explored here.

There is one final link to be drawn between the morally autonomous agent and the agent who acts responsibly: their actions (judgements) cannot be the result of irresistible coercion. If someone holds a gun to my head and demands that I drive an unsafe vehicle to Edinburgh, then I could plausibly judge one way and act another, in that I could *judge* responsibly that it would be best if I didn't, and yet be compelled through fear to act against that judgement.³⁷ This is not the result of slapdash, arbitrary or culpably ignorant reasoning - even if I decide that actually it would be best to drive to Edinburgh, I have not decided to re-evaluate my views on public safety just for the hell of it, but am being forced under obviously sub-optimal conditions to evaluate various likely outcomes for my own safety and for public safety if I accept/refuse. Consequently, it would be very odd to lay the blame for the resultant accident on my irresponsibility, and indeed relatively implausible even to suggest that I *had* acted irresponsibly.

As yet, there is no reason to take this argument as anything more than a motivation to think that our ordinary talk of responsibility carries a number of fairly distinctive senses, one of which seems to be best captured by "morally autonomous". In the next subsection, I make the further claim that this view of moral autonomy as requiring responsible reasoning and action - which is, I have suggested, the correct one - comes entirely apart from personal-political autonomy. There are sufficient conceptual differences and counter-examples for an agent to be morally autonomous without being personally-politically autonomous, and vice-versa. This being so, we have no reason to think that the functional analysis of autonomy as being powerful and authoritative over our lives will persuade us to adopt moral autonomy.

3.2 *The (Non-)Kantian Gulag and the Two Types of Autonomy*

In *Virtue & Taste* (1991), Flint Schier posed a thought-experiment he dubbed 'The Kantian Gulag' as part of an argument that autonomy was not simply equivalent to freedom. As part of the experiment, we are to imagine a good

³⁷ There is something appealing about the thought that any moral judgement is in some sense *prima facie* until we're confronted with the reality of it, but we surely don't want to take on the theoretical burden of claiming that it's impossible to make an autonomous moral evaluation without being so confronted.

Kantian who has been - for whatever reason - incarcerated in a prison camp. There are various crises or moral problems we can imagine this agent dealing with according to the Categorical Imperative; they refuse to give up fellow prisoners for torture, do not allow themselves to become ruled by motives of vengeance, and so on. Is this agent autonomous in the Kantian sense, and can we imagine such a situation without conceptual confusion? Schier affirms both, and then asks whether we would say that the agent was free because autonomous - to which the answer, as he very plausibly claims, is "no".

Returning to Schier's example, consider the morally autonomous agent not in strictly Kantian terms, but with moral autonomy conceived of as requiring responsible reasoning. Can this agent - at least potentially - meet the requirements given above? Surely so: they can recognise their continued agency, apply non-arbitrary rules and principles, and so on. Depending on other ethical commitments, we might say that actions like informing on one's fellow prisoners are immoral and irresponsible, immoral but responsible, wrong but all things considered excusable, or any of a multitude of other verdicts. The key thing is that it is certainly conceptually, and evidently historically, possible for persons to maintain a high level of moral autonomy even under conditions of extreme duress.

But precisely those conditions that make moral autonomy impressive make personal-political autonomy impossible - just as, in the original example, we would say that the agent was autonomous but not free, here we should say that they were morally autonomous but not personally-politically autonomous. An agent's considering the morally relevant features of an offered choice between snitching and starvation, rather than unthinkingly grasping at survival, satisfies in heroic fashion the requirements for responsible, and as such autonomous, moral evaluation. An agent being in a position where they are forced to choose between their moral integrity (or self-conception) and survival fails in calamitous fashion the basic external requirements for being powerful and authoritative over one's life.

To reiterate, the distinction is not that entirely different sets of psychological capacities or capabilities are used in the exercise of the two types of autonomy. The point is rather that the concepts are deployed in different situations, or have different goals. Personal-political autonomy is not impossible in a prison camp because there is some singular and very particular ability to decide on life goals which is only active when we are at liberty, but because the possibility of achieving those goals is dependent upon our being at liberty.

The wrong, furthermore, of forcing a prisoner to snitch or be starved is of a somewhat different one from the wrong of imprisoning them (or even ‘just’ starving them, with no bargain offered). Depending upon how far coercion excuses the prisoner, the wrong partly inheres in making them complicit or straightforwardly using them in some further immorality. In other words, if the prisoner experiences a failure of moral autonomy the disvalue lies in the fact that they have thus acted wrongly or immorally, not (from the perspective of assessing *their* moral status, rather than that of the captor) that they have themselves been wronged. With failures of personal-political autonomy, the wrong is the very direct harm to the agent, not that the agent is thus dragged into the commission of immorality themselves.

We can sketch many further examples where an agent can act morally autonomously without being personally-politically autonomous: disaster situations or conditions of extended chronic shortage, for example, have conceptual space for agents who are rigorously fair, non-arbitrary, reflective etc. in their moral judgements but cannot so much as decide whether to fish from one creek or the other.³⁸

Approaching the question from the other angle and asking whether an agent can be personally-politically autonomous whilst being morally heteronomous gives another affirmative answer. Take a “Ripley case”, involving someone who is pathologically incapable of seeing others as making moral demands on them.³⁹

³⁸ To give one concrete example, there is no reason to think that someone caught up in the civil war in Syria suddenly becomes incapable of responsible reasoning, but they might very well have almost no room for self-direction.

³⁹ Here, ‘Ripley’ refers to the villain protagonist of Patricia Highsmith’s books, rather than

Will they make non-arbitrary moral judgements about, for example, distribution of resources, or acknowledge the possibility of strong countervailing claims? No - indeed, Ripley *cannot* do so. But will they make and carry out courses of action according to some deeply-held principle or self-conception (even a bad or harmful one)? There seems absolutely no conceptual reason that they should not, and fairly good observational evidence to suggest that people in fact *do* lead morally heteronomous - albeit generally not to Ripley's level - but personally-politically autonomous lives.⁴⁰

Without yet delving too deeply into episodic/programmatic and local/global distinctions at this stage, it is also worth remarking that the asymmetry between moral and personal-political autonomy is mirrored at both levels - the examples so far have all been writ large and concerned with what we might paint as autonomous *lives*, but I can have a local failure of moral autonomy, such as making a decision about how to interact with someone based solely on my dislike of them, without any restrictions on my self-direction. Similarly, there is no reason to think that I must reason less fairly or rigorously because I have been temporarily prevented from going to my favoured pub. Just as in the wider-ranging cases, one is a (minor) breach of my obligations to others, and the second is a (brief) restriction of my own valued self-direction.

There is one move remaining to someone who wishes to claim that moral and personal-political autonomy serve or are served by the same function.⁴¹ It will

Sigourney Weaver's tough-as-nails space trucker. For further discussion of Ripley-type cases, see Oshana (2006), pp165-170.

40 Again foreshadowing a forthcoming argument, I would note that many of these people might live fairly personally-politically autonomous lives in spite of themselves - someone may *think* that "just being born doesn't mean you're owed anything" and be autonomous, but this is in part thanks to precisely the type of social structures they explicitly claim to reject.

41 There is also a Kantian move, which I have not considered more fully owing to constraints of space and because it is somewhat orthogonal to my task here, but which can be summarised as follows. It is not, in fact, possible to act responsibly (and thus meet the necessity condition for moral autonomy, as I claim) and wrongly, the Kantian says, because acting according to the right kinds of reasons, i.e. autonomously in the Kantian sense, *always* entails that one is acting correctly. My response, similarly summarised, is that I can remain silent on what precisely it is to make responsible moral decisions. If the Kantian asserts that anyone who is *really* deliberating suitably hard and following the correct decision-making procedure will act correctly, my claims here can accommodate that perfectly well, because all I have suggested is that responsible decision-making is equivalent to moral autonomy - I have not made any strong claims about what is required for morally responsible decision-making. For the Kantian, in other words, responsible moral reasoning cannot result in morally wrong action.

say something like this: someone who is incapable of or prevented from being apt for one kind of autonomy will necessarily be inapt for the other as well. *Contra* my claims about prisoners, an agent who is forced to make a choice between their moral integrity and their survival can be neither morally nor personally-politically autonomous, and the same conditions are doing the malevolent work in both instances.

The plausibility of this objection depends on its strength. If we think it sociologically very unlikely that an agent with extremely limited power and authority over their lives will be morally autonomous, then it seems fairly reasonable (although by no means infallible), but this is no challenge to my theory. It is difficult to think of a form of coercion or manipulation which is on the one hand potent enough to undercut an agent's reflective capacities, but on the other has no practical impact on their self-direction. We should not be surprised that many circumstances which are hostile to one form of autonomy are also hostile to the other.

If, though, the claim is interpreted as saying that someone who fails to meet any of the preconditions of one kind of autonomy must also fail to meet any of the other, then it is nearly trivial. We have already seen that many of the same capacities (reflection, some measure of self-control, instrumental reasoning etc.) are required for someone to be capable of either form of autonomy. To argue from this that moral and personal-political autonomy are irreducibly equivalent is akin to saying that football and rounders are conceptually inseparable because they are both impossible to play without physical extension: it is true but uninteresting on a coarse-grained account (which misses everything that matters), and it is false on an account fine-grained enough to capture the relevant distinctions.

This being so, it is implausible to claim that the idea of autonomy as being powerful and authoritative over one's life is equivalent to the concept given us by moral autonomy. To be morally autonomous you must indeed have some kind

But this is solely as a result of their particular theory of responsible reasoning, and so my claim that moral autonomy requires responsible moral reasoning is true after all.

of power and authority - the power to decide for yourself, and the moral authority attached to responsible reasoning - but moral autonomy requires far less control over the direction of your life, and far less external recognition of your actions as authoritative, than does personal autonomy as we have understood the concept so far. As such, the functional analysis of personal autonomy will neither require that agents act morally responsibly, nor that morally autonomous agents be powerful and authoritative over their lives - whatever the constituents of personal autonomy turn out to be, they will not be the same as those of moral autonomy.

Conclusion

In this chapter, I have made three moves. First, I argued that we should approach the concept of political authority via a functional rather than conceptual analysis: what it is for some institution to be authoritative is dependent upon what that institution's purpose or role is, and so the authority of political institutions is greatly determined by substantive rather than procedural matters. A right-libertarian and a social democrat, for example, will have different roles for political institutions and hence different success criteria for when those institutions may legitimately claim authority.

I then made a parallel argument for the concept of personal autonomy. We want such a concept to be able to account for the value we place on self-direction, and to account for those features of a life which seem important for such self-direction. The concept which does this, I suggest, is one of the autonomous agent as being powerful and authoritative over the course of their lives: able, within limits, to choose how her life goes according to what she values, and have that decision respected and (at least *prima facie*) enabled in virtue of her being a legitimate source of values. This concept captures many of our intuitions about autonomy, without being vulnerable to the kinds of objections that plague many particular *conceptions* of autonomy.

Finally, I sought to head off the idea that moral autonomy could serve the function of accounting for self-direction in the way set out above. To be morally autonomous is, I argued, essentially to be a responsible moral reasoner, and there is nothing about that status which requires the kind of wide-ranging power and authority necessary to be plausibly described as deciding how one's life goes. We can therefore safely leave conceptions of moral autonomy aside when we seek a plausible account of autonomy as power and authority.

In the next chapter, I turn to constructing a conception of personal or personal-political autonomy which holds that being powerful and authoritative over one's life is a matter of standing in particular social relations.

3.

The Structure of Autonomy

Introduction

The preceding chapters have argued that the concept of autonomy is not in principle objectionably individualistic, and that a properly-conceived account will show autonomy to be consistent with a thorough-going commitment to social values. The next step, obviously, is to provide such an account.

Examining the structure of personal-political autonomy entails two tasks: identifying the components or features relevant to being autonomous, and identifying how they interact.¹ The most minimal accounts obviously have an easier time of this. For instance, Frankfurt's (1971) structuralism has essentially two components, lower and higher-order desires, which interact fairly straightforwardly. If our higher-order desires endorse our lower-order desires, we are autonomous. Frankfurtian structuralism is a paradigm example of what I call "internalist" accounts of autonomy: those features relevant to our autonomous status are all internal to our psychology.

In contemporary debates, it is more usual to adopt some flavour of "externalism" about autonomy, i.e. to argue that at least one of the determinants of our autonomy is external to our psychological states. Procedural accounts are externalist in this way: we must consider the procedures through which first-order desires were formed, and the factors which affected our decision-making and psychology, as well as our own higher-order attitudes to the desires (and how *those* attitudes were formed). Desire-satisfaction theories are on the face of it the simplest form of externalism: the crudest possible conception would have it that the more of your desires are satisfied, the more autonomous you are. More plausible versions, though, are like structuralist or procedural accounts with an extra condition bolted on: now we have desires; our higher-order attitudes towards those desires; how the desires and our attitudes towards them come to be formed; and lastly, whether the 'correct' desires,

¹ Hereafter, "autonomy" should be taken to refer only to personal-political autonomy unless specifically stated otherwise.

those formed under conditions of procedural independence are being fulfilled.

While each of these accounts pinpoints, to a greater or lesser degree, something important about autonomy, I claim that each fails to capture at least one feature which is necessary for the concept of autonomy to do the normative work we want it to. The functional analysis in the previous chapter suggests that in order for autonomy to account for the value we place on self-direction and those features of life which are important for self-direction, we ought to think of the autonomous agent as being powerful and authoritative over her life. Here, I make the further argument that no internalist or desire-satisfaction externalist conceptions give us the proper criteria for being powerful and authoritative, and so we must adopt a still more stringently externalist view.

On my view, autonomy has internal preconditions but actually being autonomous is a matter of external, non-psychological features of our lives. Most distinctively, I claim that autonomy is constituted largely by our standing in particular social relations. In many cases, these relations seem to be entirely external to and independent of any desires we have: we can be autonomous and unhappy with our decisions, or autonomous “in spite of ourselves”, so long as we stand in relations which constitute our being powerful and authoritative over our lives.

Complexity is not necessarily a virtue, of course, but it seems to me to be the only way to satisfactorily account for the internal preconditions and external conditions for one’s being powerful and authoritative over one’s life - i.e. being autonomous - which I’ve so far endorsed. As such, the complex nature of the theory is a necessary consequence of its attempt to elucidate a key value of our generally pretty fiddly lives.

The aim of this chapter, then, is to work out what the key structural features of autonomy are and how they fit together. I show how the relational account fulfils the requirements laid down for autonomy: an agent is autonomous insofar as she is powerful and authoritative over her life, and this status is constituted by her satisfying certain psychological preconditions and standing in certain relations. The most pressing task is therefore to explain how (and why) the relational account best meets the power and authority criteria of autonomy.

In order to accomplish this, we must first be clear about our reasons for thinking that the criteria for autonomy *are* power and authority; and, keeping in mind the worries raised in the second chapter, show that being powerful and authoritative is not a chimerical or individualist value.

1. The Concept of Autonomy

Assessing the plausibility of a particular conception of autonomy, of course, first requires that we know what the concept is supposed to do - you can hardly decide whether or not to buy a particular car without having a reasonably good idea of what function a car is supposed to serve. Similarly, if you know what a car does, but are unable to explain why that should be important to you, it would seem odd to assert that you should buy this, that, or *any* car. The aims of this section are thus to lay out a fairly uncontroversial sketch of the concept of autonomy, suggesting that at the broadest level autonomy consists in having power and authority over oneself, and to give some more detailed justifications for the presumption that autonomy is a valuable thing (a presumption which, as the criticisms raised in Chapter 2 show, is by no means self-evidently true) by analysing the concept in functional terms.

It is worth re-emphasising the methodological approach adopted in Chapter 2 before continuing. Autonomy clearly serves a normative purpose - we refer to it in discussions of redistribution, political power and medical ethics among many others - and so we must be clear about what that purpose is, and ensure that the concept and conception we offer is fit for purpose. Taking the approach that such-and-such a conception provides plausible necessity and sufficiency conditions for our pre-theoretic intuitions about autonomy, that the conception would allow (say) voluntary slavery, and that therefore one can be an autonomous slave, is to misapprehend the role that the concept of autonomy is supposed to play. Instead, we should assess accounts of autonomy by examining how well they perform the desired function of explaining and illuminating the value and practice of self-direction. I take my task here to be the elucidation of a concept of autonomy that best serves the end of valuing agency, rather than

strict analysis of a pre-existing concept in an attempt to discover its necessary and sufficient conditions.

What Does Autonomy Require?

To begin by outlining again the concept of autonomy: something shared across almost every contemporary or near-contemporary account is the idea that self-determination is key.² Self-determination (or self-direction) can remain fairly broadly defined for the moment, taken to mean something like the ability to exercise some non-trivial level of control over our lives, being able to make “choices and [take] actions that effectively accord with deeper wants and desires”.³ One familiar set of contrast cases which can illustrate the concept, as well as distinguish it from theoretical neighbours, is Pettit’s (1996) case of the citizen and the slave. By considering the contrasting situations of the agents, we can identify the differences which seem to make one autonomous and the other not (or less) autonomous, and then move to explain those differences.

One concept which is often conflated with autonomy is that of autarchy, simple unrestricted choice - I am autarchic insofar as I choose what I want and do what I want (the negatively free agent is paradigmatically autarchic).⁴ It will be immediately obvious that neither the slave nor the citizen is autarchic, and so we have another reason to think that free choice, at least, cannot be *all* there is to autonomy. Quite how important autarchy is must remain a question to be dealt with later; for now, we can say that the slave or serf is considerably less autonomous than the citizen with rights of self-expression, freedom of movement and so on.

Insofar as the slave’s life is directed in the interests and at the will of others, they are non-autonomous. The citizen - whilst by no means to be necessarily regarded as ‘fully’ autonomous, as will become clear - has considerably more room to make her life go the way *she* sees fit; within certain not-inconsiderable bounds, she sets and pursues goals according to her own principles.

2 See, for example, Colburn (2010); Dworkin (1988); Friedman (2003); Meyers (1989); Oshana (2006); Raz (1986).

3 Friedman (2003), p28.

4 See Benn (1976) and Colburn (op. cit.) for more on the distinction between autarchy and autonomy.

What, then, are the differences that explain their contrasting states? Obviously the slave's enslavement is crucial to their lack of autonomy, but to explain that a slave is not autonomous because they are a slave is (perhaps not quite) to argue in a circle. To put it another way - what features does the citizen possess that the slave does not?

It cannot be what we might think of as basic capacities which are required to be capable of self-direction in the first place, such as a reasonably coherent psychological identity, some amount of linguistic ability, instrumental rationality or the like. There is, obviously, no necessary difference between the slave and the citizen in this respect: both of them could be equally well (or badly) *equipped* for autonomy without being equally autonomous.

The differences instead lie in the extent to which they can actually use these capacities, and the extent to which this usage follows from or is consistent with the agent's demand for recognition.

The first difference is comparatively straightforward, and can be seen in the example above: A and B both have identical capacities, however these are understood for autonomy, but A has the freedom to employ those capacities whereas B does not. A is therefore more autonomous.

It is worth taking a moment here to set up the second difference, that of recognition respect.⁵ This demand for recognition is not something like a demand for the attribution of professional excellence or interpersonal importance, although these might fall under it. Rather, it is a more basic recognition of oneself or another *as an agent*. When we recognise others as agents, we recognise - or presume - their possession of normatively-important capacities like a sense of self and some measure of instrumental reasoning. Arguments about what, precisely, are the conditions for agenthood can be laid aside. Whatever we think makes an agent an agent, we think it possessed by those we recognise in this relevant way, and there is one consequence of recognition which is important to my argument.

5 Darwall (2006) expands on the normative role of recognition respect, and is the primary influence on the (much-simplified) account of recognition presented here and in later chapters.

When we recognise somebody as an agent, we grant their wishes *prima facie* authority; my being an agent means that there is no general requirement for me to explain or justify my actions unless certain other conditions (such as my actions impinging on others, to give one example) apply. “Because I want to” is, *ceteris paribus*, all the motivation anyone can feel entitled to get out of me, and all the justification I need to perform the action.⁶

This *ceteris paribus* is, nonetheless, doing a job of work here - or else my account *would* simply collapse into autarchy-as-autonomy. If I want to go round setting fire to orphanages, then all else is not equal - my shrug of “I just wanted to” does not defeat moral arguments against immolating children. Note, though, that we still have to have a non-arbitrary reason to rule out “because I wanted to”. The presumption is that we may decide how best to direct our lives; when we are permitted or supported in a particular action in virtue of it being what we want to do, we are recognised as agents and we are authoritative over our lives. We might be forbidden from doing some things without this indicating a lack of recognition respect, but there must be good reasons for this interference.

To recap, and return to the contrast case of the citizen and the slave: the citizen is able to self-direct in a non-trivial manner, and her self-direction is regarded as at least *prima facie* valuable just in virtue of her valuing of it. The slave or serf, conversely, has very limited self-direction, perhaps none at all that isn't trivial, and their wanting to do something is not acknowledged as a reason for them to do it. The key distinctions between the cases, I will show, are distinctions of power and authority. The autonomous agent is powerful and authoritative with regards to the direction of their life, while the heteronomous agent is not (this is a scalar rather than binary distinction - we are more and less powerful and authoritative, not “godlike” and “totally powerless”).

⁶ It's worth being crystal clear that this isn't intended to suggest that being an agent entails that we can never be legitimately asked for, nor required to give, more justification than this: if, for example, I try and evade childcare duties by shrugging and saying “I didn't want to”, this response will be rightly regarded as unsatisfactory. Rather, the claim that we're not in general required to give more justification than this is supposed to capture the thought that being an agent entails that unless I'd breach some obligation, or fail to fulfil some duty, or my behaviour is out of character (and so on), merely wanting to do something is a good enough justification for us to do it.

While the citizen can make at least some use of these basic capacities or preconditions - she can decide that a day off at the end of the week is worth working late on Monday and Tuesday - the slave has a far more limited arena to exercise self-direction, not to mention a more restricted number of options even within that arena. They are, to reverse Raz's aphorism, free to decide whether to cut down trees in this field or the next, but not free to give up being a lumberjack. In short, the citizen has considerably more *power* over their life than the slave does. The autonomous agent must, then, both have the ability and the regular opportunity to self-direct; "power" here should be understood as something exercised instead of a mere possession of the appropriate mental and physical powers, whatever they might be, for the exercise of autonomy.⁷

A similar point can be made regarding the *authority* of the agent. For the moment, this authority should be thought of simply as a kind of normative recognition as described above; a status such that having a particular desire is regarded, simply because of one's being an agent, as a *prima facie* fairly strong reason for one to fulfil it (just as being recognised as an 'authority' on geese means that your saying something about geese's behaviour gives others *prima facie* reasons to believe it). This *prima facie* reason is obviously defeasible: if I am blind drunk or suffering some other serious impediment to (say) my instrumental rationality, then the *prima facie* reason given by my agency does not end up being a good reason all things considered. But I take it that this is part of why it's only a *prima facie* reason in the first place, albeit a weighty one: we take agents' desiring to do something as being on the face of it a good motivation for them to do it (and perhaps for us to assist), because we presume their rational and epistemic competence. But if it then transpires that the agent is somehow incompetent, the motivation ceases to be a good one. If our avian authority asserts "geese display such-and-such behaviour", we believe them in the normal run of things - their assertion is a good *prima facie* reason to think that geese do indeed behave like this - but if they then assert something like

⁷ The autonomous agent could still very well decide that (originally non-voluntary) social commitments are a valuable part of their identity, and need not be constantly engaged in self-examination or grand projects. To say that self-direction must be exercised is simply to say that an agent cannot have their life and values wholly set by another and be counted as autonomous, no matter how much "potential" for autonomy they have. A potential chicken is still just an egg, after all.

“geese are small rodents which enjoy peanut butter”, that *prima facie* reason is defeated.

To return to the citizen and slave case with respect to authority: the slave may have some discretion regarding certain tasks or some parts of their life, but this limited power to self-direct is not based on the same foundations of recognition as that of the citizen’s power. To a greater or lesser extent, the power to self-direct is regarded as default for a citizen - assuming that some course of action is not ruled out by morality, then the citizen should regard herself as being entitled to pursue it just because she wants to, or because it leads to some other outcome she values, etc. In short, the citizen’s power over her life is matched by authority in a way that the slave’s is not; when the slave is left to choose whether to repair *this* end or *that* end of their exploiter’s wall, the choice does not follow from recognition of their status as an agent capable of making normative claims. They express no self-conception in the choosing, and they are not offered the choice in virtue of their agency.

Consider: we do not allow Roombas free roam because we think that they are independent sources of value; we leave them to potter about because it is of no particular importance to us which corner of the carpet it cleans first. Similarly, the slave’s being left to choose between ends of the wall does not constitute their being granted room for self-direction on the basis of recognition of their authority over their life (indeed, such a bequest would seem conceptually at odds with slave-holding in the first place), but a mere absence of active interference.

This is not to suggest, of course, that an enslaved agent ceases to be, or to see themselves, as a source of value, nor that they are prevented from *all* self-direction which they hold to be important. Rather, there is a sharp disconnect between what they can do, and what they can do on the basis simply of being an agent. Where the citizen’s authority is at least theoretically the default position, the slave defaults to having no authority - their realm of self-direction is set according to the wishes or values of another.

The authority criterion given above shares a certain amount of conceptual overlap with ideas that self-respect is necessary for autonomy - one cannot, the

argument runs, be an autonomous agent without believing that the goals one has are valuable, or that one is capable of setting and following through on goals.⁸ Given this overlap, the differences are worth drawing out by way of explaining why authority here is not simply reducible to self-respect.

Since it seems possible that an agent could engage in meaningful self-direction without realising it (perhaps they believe themselves to merely be acting out preset patterns), or without believing themselves worthy of it, we have a good indication that power on its own is not sufficient for autonomy. What seems to be missing from the deterministic and self-abnegating agents is the kind of self-respect that allows one to hold that “because I want to” is a good (though not indefeasible) reason to carry out an action. However, the weak authority requirement is not *simply* a self-respect criterion. As aforementioned, an enslaved agent could well and correctly view themselves as an independent source of values - but still, their justification or reasoning for performing many or most actions is not going to revolve around this view.⁹ That is to say, someone could easily hold that they deserve to act from their own values, and also recognise that many or most of their actions are not of this kind.

In summary, the weak authority condition (hereafter WAC) is composed of two individually necessary and jointly sufficient sub-conditions: the agent must have some minimum level of self-respect for themselves as an agent, and the actions they take must, in general, be explicable or defensible just by reference to recognition of this agency. Taken together, satisfying these conditions constitutes being weakly authoritative over one’s life.

I have argued so far that the concept of autonomy requires power and authority over our lives, such that we can (to an extent that falls short of total self-sufficiency or implausible freedom from social conditions) manage the direction of our lives in accordance with our principles or values. Such a concept is not inherently harmful, and nor is the pursuit of it unduly likely to bring harm to the agent or to others around them (although it *may*, and may do so in a morally-

⁸ See, for example, Hill (1991).

⁹ I call it a “weak” authority requirement because it does not, as yet, require that anyone *else* acknowledge the authority of the agent. I return to this limitation in the ‘Relational Autonomy’ part of Section 3.

blameless fashion). So, we at least seem to have defeated one of the challenges to autonomy raised in Chapter 2.

This nevertheless leaves a great deal of work. Even if it is not *necessarily* the case that power and authority are conceptually or practically harmful, we have already seen that it can still be the case that particular conceptions of them are. If it turns out that in fact, the only consistent conception of autonomy involves (say) power as unrestricted choice, then the previously raised objections that autonomy is an overly individualistic or androcentric value would be difficult to refute. The task for the remainder of this chapter, then, is to show that autonomy properly conceived is consistent and not vulnerable to such criticisms.

After criticising several relatively popular conceptions of autonomy - the internalist, structuralist view most often associated with Frankfurt, an account of autonomy as procedural independence, and what I snappily title the procedural independence desire-satisfaction view - I make the case for my own account, a relational conception which owes much to Marina Oshana. On this conception, autonomy is constituted, rather than merely causally affected, by our social relations: it is not that I become less autonomous in virtue of my social status causing some desire to be frustrated, for example, but because I stand in relations which constitute my being (again, more or less) powerless to direct my life.

This account, I argue, best meets the power and authority conditions which are required to explain autonomy's role and value. There are internal psychological preconditions that we must satisfy in order to engage in the relevant kind of self-direction, and external, social relations which determine the extent to which we actually engage in such self-direction. To be powerful is to be able to pursue (morally unobjectionable) goals, to be authoritative is to have a normative status as a proper source of value for those goals; and the relations we stand in are constitutive of being either.

It is primarily the relational aspect which gives my account a reply to worries about ethical and structural individualism; to value autonomy consistently on the relational account, one must value the autonomy of others, and such an equality

militates both conceptually and in terms of practical organisation against a world organised for the purposes of reclusive mountain men.

2. Internalist Accounts

In this section, I rehearse and criticise two accounts that can be broadly classified as internalist - that is to say, they cast autonomy entirely or primarily as a matter of our psychologies or mental states lining up in the correct fashion. Only structuralism can strictly be painted as a purely internalist account; procedural independence takes some heed of external factors. I categorise proceduralism as an internalist view because external factors are of secondary importance to the proceduralist thesis - so long as no-one is malignly interfering, the key concern is whether we meet the various conditions of competence (such as instrumental reasoning and so on), and these preconditions relate to our psychologies, not our surroundings.¹⁰

To take the accounts in order, I now turn to structuralism.

2.1 Structuralism

Structuralism's great virtue is the straightforwardness of its conditions. It is entirely conceptually possible for any given agent on Earth at this moment to be autonomous if we adopt, for example, Frankfurt's view: autonomy consists in having first-order desires which are identified with or endorsed at a second-order level, and being motivated to act by these desires.¹¹ If I desire (to do) something ("I want to eat that cake!"), then I have a first-order desire; if I have a desire for or positive attitude towards that first-order desire, I have a second-order desire endorsing it.

¹⁰ This is probably a somewhat contentious categorisation, but ultimately whether or not one agrees with my taxonomy of proceduralism is a side-issue to the point at hand - I would be happy enough calling it semi-internalist, or pseudo externalist, or the like.

¹¹ One version of this view - although he does not refer to it as an account of autonomy - is given by Frankfurt's 1971 paper 'Freedom of the Will and the Concept of a Person'. See also Dworkin (1976); Dworkin's version is often referred to as the identification or split-level view, but for simplicity's sake (and because there is more overlap than not) I refer to them interchangeably as structuralism.

If, for instance, I have some set of (first-order) desires $A_1 \dots Z_1$, all of which are endorsed by the second-order set $A_2 \dots Z_2$, then on the structuralist account, I am autonomous regardless of what kind of desire A_1 (or A_2) may be, and autonomous regardless of whether I will ever fulfil this desire.¹² I may, then, have a desire to go and enjoy the countryside, and after checking this desire against the various higher-order desires I have, I find that I can endorse it - that is, I'm not riddled with guilt over wasting valuable research time, nor secretly convinced that hillwalking is a bourgeois conceit. I identify myself "*decisively* [italics Frankfurt's] with one of [my] first-order desires", such that I am happy to be moved by it, and so the structuralist will claim that I am autonomous.¹³

Something important about autonomy is obviously captured by this kind of account. Without returning to the outdated notion of autonomy as being the self-mastery of the rugged individualist, we can still hold that someone who is consistently acting on desires they regret or repudiate (or, conversely, consistently fails to be motivated by desires/attitudes they regard as important or praiseworthy) is alienated from themselves in such a way as to render them non-autonomous. They may still have power over their life, since power over one's *desires* is a very different thing from power over one's actions: I can on the whole stop myself from eating cake much more easily than I can stop myself from desiring cake. However, they plausibly fail the weak authority condition by dint of their actions not reflecting their deep values and commitments. One of the platitudes of autonomy, after all, is that it is valuable - or, at least, we value it - because leading a life "of one's own" is valuable, and we can hardly be

12 The 1:1 relation between first and second-order desires here is assumed purely for simplicity's sake, and should not be taken to imply that structuralist accounts are committed to a requirement that every first-order desire must have a second-order desire devoted solely to endorsing it in order for the agent to be autonomous. Such an account would, it seems to me, require an implausible level of self-consciousness about everyday actions and modes of life or else label such habits heteronomous: introspection reveals that, despite a marked preference for some kinds of soft drink over another, I have no discernible second-order desires whatsoever about which colour of fizz I ought to drink. Certainly, there is no *decisive* identification; but does this make me less authoritative or powerful over my life? I cannot see how, unless, again, to be powerful and authoritative (i.e. autonomous) is simply to be some kind of ultimately autarchic, rigorously self-creating Great Will - and this is a conception of autonomy that has already been rejected. Bratman's more recent claim that "autonomy is realised by the functioning of reasoning-guiding policies that are higher-order" (2003, p171) evades the particular worry of unduly rigorous self-consciousness which I have rehearsed here, but as a structuralist theory is still vulnerable to the Seamus-type problem cases I set up in this chapter.

13 Frankfurt (ibid), p16.

leading a life of our own if all of the desires which partly define us are desires that we would not, or would rather not, recognise as ours. In short, something like a more relaxed version Frankfurt's freedom of the will looks as if it is probably necessary for *authenticity* - and, to a lesser extent, the authority condition of autonomy.¹⁴ As I now argue, however, it is not sufficient for autonomy.

2.2 *Seamus, Sinead and the Problems with Structuralism*

Just as it would be odd to think of the autonomous agent as one with no reflective capacity or control, it would be very unusual to say that so long as I was at peace with my desire to stand atop a hill, whether or not I can actually do so is irrelevant to my autonomy. Consider a close possible world where my equivalent, Seamus, has the same deep desire to roam o'er hill and glen, but unlike me, is being held prisoner for an indefinite period. How, we might think, can it possibly be the case that I and my counterfactual compatriot are equally autonomous despite my being free to fulfil the desire, and his being imprisoned and unable to fulfil it? Here, Seamus patently has very little power over his life, and while there are doubtless significant ways he can act from his own values (refusing to speak to screws, hunger striking, or whatever), the vast majority of his self-direction will be of the "walking in this corner of the prison yard rather than that one" sort, and as such not explicable by reference to his commitments or values.

To cast Seamus as autonomous when we have defined autonomy as being powerful and authoritative thus seems inappropriate. The freedom to walk in one direction instead of another counts for little if we are only picking between the directions by virtue of being imprisoned, after all. Since this worry afflicts other internalist accounts that can evade the structuralist-specific issue below, I temporarily set it to one side.

A further and familiar problem awaits structuralism. In the case above, it seems that Seamus is not autonomous - but it does at least seem that insofar as his

¹⁴ Although the extent to which our desires must be authentic in order to be properly counted as part of our identity is very much up for debate - for present purposes, I assume that some measure of authenticity is necessary, but make no claims regarding how much or how precisely the authenticity of a desire is to be specified.

desires are “in order”, he is more autonomous than he otherwise would be. The problem, in other words, has been that the states identified by structuralism are insufficient for autonomy, not that they are inconsistent with it. As has been famously demonstrated, however, inconsistency does arise in at least some cases.¹⁵ Imagine Sinead, a woman similar to Woolf’s “angel in the house” - conforming to 1950s gender roles, subservient to her husband in the economic and domestic, and perhaps sexual, domains; and furthermore, happy with such a life. Sinead values the role of home-maker, and actively disvalues things like economic independence or a career; and her higher-order reflections on these attitudes lead her to endorse them. She identifies decisively with her lower-order desires and attitudes, and on the structuralist account she is thus autonomous.¹⁶

There is an obvious criticism to be made here regarding the possibility of autonomy being determined by factors outside the agent’s psychology, but as it leads into procedural independence (and, ultimately, thorough-going externalism), it makes sense to first examine the type of objection to structuralism raised by Hill (1973) and Westlund (2003). Each has its own emphasis, but the overall outline of the objections is roughly the same. Imagine as before an “angel in the house”, but now add that for whatever particular reason, she has begun to chafe against her disempowerment. She grows to loathe doing dishes, preparing dinner, cleaning the house, and so on - but on reflection, she repudiates these desires. Despite loathing the daily prospect of getting out of bed early to make breakfast, she believes that she should *not* loathe it; what she desires is to want to discharge these various burdens. She has become, at least by structuralist lights, non-autonomous.

The problem arises here: according to the structuralist schema, Sinead will become autonomous by effectively curbing those desires which she rejects on reflection - in other words, those desires to smash the plates rather than clean them, and so on. It is only when she accepts her domination, in short, that she will be autonomous, and this is the second damning criticism against

¹⁵ Particularly in Friedman (1986).

¹⁶ I borrow the particular title and set-up for this example from Oshana (op. cit., pp58-60) although (obviously) indirectly from Woolf (1966).

structuralism: it not only strongly appears to sanction oppressive norms so long as they are identified with, but when these norms are resisted, then rather than supporting self-determination (something that even other unsatisfactory conceptions of autonomy do, albeit perhaps in unhelpful or merely formal ways), it actively urges the suppression of desires to resist, *for the sake of autonomy*.

One cannot be said to be authoritative if the first-order desires one has are repudiated for no other reason than that they conflict with higher-order desires which themselves disbar meaningful self-determination. If I have some high-level commitment to subservience, and find myself riddled with first-order desires towards rebellion, it is entirely ass-backwards to say that my autonomy, my ability to self-direct, is best served by quashing these selfsame urges to self-direct. As Friedman says with regards to an agent playing the angel in the house role: “Her frustration, grief, and depression, and the motivations to change her life which spring from these sources, may be her only reliable guides”.¹⁷

There is, unsurprisingly, more than one interpretation of what has gone wrong in cases such as these. Friedman casts the problem as one of psychological or reflective misfires caused by oppressive socialisation, while (for example) Oshana plausibly argues that whether or not the agent’s reflective capacities have been harmed, she is still not autonomous.¹⁸ This distinction will become more relevant in the next section, but for the moment, I remain silent on the matter. The important claim at this point is just that it cannot be the case that having all your psychological ducks lined up in a row is enough to make you autonomous.

Despite the force of the claim that we must have (at least some) ownership of our desires and goals - we cannot be metaphorical automata furrowing someone else’s groove, nor alienated from ourselves to the point of psychological fracture - the implausibility of freedom of the will as a sufficient condition for personal-political autonomy suggests that external factors must play a significant role in

¹⁷ Friedman (1986), p31. Interestingly, Friedman doesn’t initially seem to hold the angel in the house objection as doing much work in criticising structuralism; the majority of her paper is focussed instead on pressing home a charge of infinite regress.

¹⁸ Oshana (2006), pp59-60.

determining our autonomous status. Autonomy is not simply a matter of well-ordered desires: its structure extends beyond the self.

2.3 Procedural Independence

Worries of the kind laid out above have caused theorists such as Christman (1991, 1993), Gerald Dworkin (1988) and Meyers (1989) to pay greater heed to the status and history of the agent, whilst still putting the locus of autonomy firmly within their psychology.¹⁹ The obvious weakness, they would argue, is that autonomy conceived as bringing first-order desires in line with higher-order desires, or of having authentic first-order desires, must also presume that these higher-order desires are themselves authentic - and in the Sinead case, it seems overwhelmingly likely that the relevant higher-order attitudes are the result of endemic social pressures which are hostile not only to her self-determination, but also to the critical analysis of these social pressures. The procedures by which her desires have been formed are themselves moulded and infiltrated by factors inconsistent with her “deciding for herself”, and hence she is not procedurally independent.

That is to say, Sinead’s social/material conditions push her towards a particular role and set of attitudes, and that role and/or those attitudes themselves, as well as the social and material conditions which they are a product of, militate against critical reflection. To use a familiar phrase, Sinead has internalised exactly those norms which oppress her.

It is thus fairly easy to see why higher-order repudiation or endorsement of first-order attitudes fails to act as a suitable criterion for autonomy. It is as if, being effectively convinced that punching oneself in the face is good for one’s health, I regard my own disinclination to do it as signifying a failure to be sufficiently concerned with my physical well-being. The problem is not that my health is an inappropriate thing to be concerned about, but that there are certain epistemic preconditions I have to satisfy in order to effectively *be* concerned with well-

¹⁹ Although generally the agent’s history is important only insofar as it has epistemic effects: they are “historical” accounts of autonomy in the sense that one’s history may determine one’s current autonomy.

being, or for my concern to be directed in the right kind of way to achieve its goal.

The Dworkinian model of procedural independence does fairly little to extend the scope of autonomy. Rather than higher-order identification with desires, he suggests that what is necessary is instead that the agent be able to reflect on their desires and change them in light of these high-level attitudes.²⁰ This capability need not be of the wearingly self-conscious kind of reflection typically engaged in by philosophers; someone can equally as well display procedural independence by deciding to give up smoking because of a higher-order desire for children, or not attending works events they find objectionable, as by sitting in their armchair until they have concluded that publishing in some journal really does cohere with their theoretical political commitments. Indeed, they can be procedurally independent without engaging in any behaviour or attitude modification at all, so long as it would be possible for them to engage in such modification if they wanted to.

The history of the agent now becomes important. If they have been socialised in such a fashion as to harm their ability to critically reflect - if, for example, they have been taught since a young age that critical thought is a mark of disrespect or truculence or the like - then their higher-order identification with their first-order desires cannot be taken as signifying autonomy.

To illustrate the point, imagine someone who is misinformed that some particular brand of fizzy juice contains dangerous levels of carcinogens. Further, they are tricked into thinking that the company which makes the purportedly offending drink are engaged in some spectacular guerrilla PR. As far as the agent knows, they have paid off everyone from high-profile scientists to newspaper editors to folk on the metaphorical Clapham omnibus to claim that the drink isn't dangerous at all, but instead a delicious low-calorie refreshment. In such a case, the agent will have a desire not to consume the drink, and indeed will have a higher-order identification with this attitude of distaste, but it would seem misguided to say that they are (with respect to this matter) acting

²⁰ "Autonomy is conceived of as a second-order capacity of persons to reflect 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" (1988, p20).

autonomously. After all, not only do they have a mistaken idea about the object of this attitude, but there is no way - at least, as the thought-experiment has been specified - for them to correct this mistake. Any evidence which would enable them to reflect accurately on whether or not drinking this particular fizzy juice is the sort of thing they want to be doing is, by virtue of the failure of procedural independence, going to be disbarred from their consideration.

By contrast, if they are in full awareness of the facts about the drink, and come to some eventual decision about the attitudes they hold towards it - that is, "I desire/do not desire it" and "I identify/do not identify with this desire" - then, the proceduralist will say, they are autonomous. Their desires and attitudes are either in some sense "their own" or at the very least attitudes they are happy to take ownership of, and we have no reason to think that they are taking ownership under conditions of psychological or epistemic pressure. As we might think of it, they're not buying any attitudinal pigs in pokes.

Proceduralism thus purports to satisfy the twin requirements of autonomy in the following ways:

1. *Power*: The procedurally independent agent has power over their life insofar as they act on desires/attitudes which they either have endorsed, or could endorse, after critical reflection. If I am pursuing goals that I don't want to, or wouldn't want to if given appropriate information about them, then I am still self-directing in the minimal, "this tree rather than that tree" sense. But there is none of the kind of self-direction which we generally take to characterise an autonomous life, where attitudes and the objects of those attitudes are reflective of an agent's deeply-held commitments and values. Attitudes formed or endorsed after critical reflection under conditions of procedural independence, conversely, can be considered as contributing to or constituting the agent's ability to shape their life in the way they would like (or in a way consistent with their commitments).

2. *Authority*: An agent who satisfies the conditions of procedural independence is authoritative in the sense that there is no reason to think their desires fail to properly or appropriately express their commitments. Imagine that you have some desire for entertainment, and decide to watch a film. You are a fan of

comedy, and your trusted friend recommends that you see *Shaun of the Dead*. Here, unlike in the energy drink case, your decision-making process has not been manipulated - your friend honestly opines that it's a good film they think you would enjoy, and that is the extent of their input - and, in the limited matter of whether or not to watch that film, you are authoritative. There is no reason to think that a decision to watch it conflicts or is in any way inconsistent with your other beliefs, desires and so on; when you claim that "On the basis of being a comedy fan I want to watch *Shaun of the Dead*" there is no epistemic disconnect in the way that there would be if we had been misinformed such that we instead claimed "on the basis of being a comedy fan I want to watch *Wuthering Heights*".

Analogously, when a procedurally-independent agent says "On the basis of some principle or desire, I want to x", we can understand her x-ing by reference to what she desires or values, and what kind of thing x is. If she had no self-respect, it would seem at least odd for her to say "I wish to x, and that wish seems like a good enough reason to x", and if she had been manipulated, misinformed, or just forced into x-ing, we would not be able to make the connection between her wishes and her action - it would be akin to trying to work out why a comedy fan had rented *Wuthering Heights*.²¹

Sinead-type cases of internalised oppression, then, can be appropriately explained by procedural independence theories of autonomy. Sinead's second-order identification with her desires to cook, clean and so on do not make her autonomous because the identification is itself the result of malevolent attitude-forming procedures; her preferences are, to use the jargon, adaptively formed.²²

2.4 Seamus (though not Sinead) and the Problems with Procedural Independence

Although procedural independence gives a good explanation of the Sinead case, the account is still problematic. Recall that Seamus, the subject of our other thought experiment, has deep desires to go hillwalking and - we can now make

²¹ This is obviously a somewhat naive picture, since we would in reality likely conclude that there must be some motivation or influence we're not privy to, but the point remains - her actions simply would not be explicable by reference to her agency or self-conception, and if that was all an observer had access to, then it doesn't seem unreasonable to suggest that they would be stumped.

²² Elster (1982, 1997) gives a helpful overview of adaptive preference formation as well as the problems it raises for autonomy theorists.

explicit - these desires have been formed under conditions of procedural independence. He identifies with them not because he has been brainwashed or manipulated into identification but because he honestly (and correctly) believes that hillwalking forms or expresses some important aspect of his self-direction. Structuralists and procedural independence theorists will both agree that Seamus is autonomous - and will both run into the obvious objection that imprisonment and autonomy do not seem like natural playmates.

Take the authority condition first. Seamus is clearly still authoritative in the same sense that a comedy fan is authoritative in their choice of film; that is, epistemically and with regard to his desires, Seamus is the authority. The screws could not claim that they knew better than Seamus what he felt about hillwalking. But this is not the whole of the authority condition: his being epistemically reliable about his desires is not taken as having any import to the way he is treated.

That is to say, while desires based on mistaken premises *might* be regarded as undermining the authority of the agent, acknowledging that the agent knows what they're talking about is not the same as acknowledging that they should be allowed to match deed to word.²³ To restate the point: that an agent's desires are held to be (more or less) authentic does not equate to that agent being authoritative with regards to them. Seamus knows that he wants to go hillwalking, but also knows that this desire has no impact on the course of his life - he cannot regard his day-to-day self-direction as meaningfully following from his own values, and cannot thus meet the authority condition of autonomy.

The problem with satisfying the power criterion should be fairly obvious and does not need much elaboration. Although Seamus is able to engage in some limited episodic self-direction - he can walk to this corner of the exercise yard rather than that, sit at this table rather than that - he clearly is not able to shape the general direction of his life according to his principles. Whatever goals and plans he has for his life, their fulfilment lies well beyond his grasp for, at

²³ In many if not most cases, of course, the difficulty in mapping out how desires have come to be formed means that we should presume epistemic authority rather than require provenance for it. Oshana (2006, pp38-39) makes a similar point.

the very least, the duration of his imprisonment, and he cannot be said to satisfy the criterion of power to self-direct.²⁴

In short, and entirely unsurprisingly, Seamus the political prisoner is not autonomous - and that procedural independence would deem him so militates against adopting that conception of autonomy.

3. Externalist Accounts

The common theme running through most of the criticisms of structuralist and procedural independence accounts is that, by dint of being entirely or primarily focussed on the psychology of the agent - the “inner citadel”, in Berlin’s (1969) phrasing - rather than their political and material conditions, they ignore obvious external hindrances to self-direction. Still, the idea that some psychological conditions are *necessary* for autonomy is a fairly reasonable one.

The obvious move, then, is not to entirely abandon psychological competence or independence conditions, but to seek out other conditions which relate to this external aspect of autonomy. In this section, I consider two such externalist conceptions: procedurally-independent desire satisfaction, which evades many of the problems of internalist accounts but ultimately retreats to what amounts to an inner citadel, and relational (or social-relational) autonomy. This latter, I suggest, is the most promising attempt to satisfy the power and authority conditions for autonomy, although there are several worries it must first dispose of.

3.1 *Procedurally-Independent Desire Satisfaction*

Recall that the most serious objection to procedural independence seemed to be that it gave the wrong result in cases where an agent’s psychology wasn’t hindered, but their ability to follow through on their desires was. The most straightforward route out of the problem is thus to take the claim “the

²⁴ The distinction between the two types of self-direction here - sitting in this corner rather than that as opposed to deciding to take up hill-walking at the weekends - maps onto the distinction between episodic and programmatic autonomy, which will be discussed in more detail later in the chapter.

autonomous agent reflectively endorses their desires” and append the extra condition “...and is capable of fulfilling these desires”.²⁵ There is now a non-arbitrary and intuitively plausible distinction between mine and Seamus’ autonomous status, namely that his goals are consistently frustrated. We have equally authentic desires, and the capacity to reflect on and change these desires and our attitudes towards them; but insofar as I can achieve mine and Seamus cannot, I have greater autonomy.

Drawing the example more broadly, imagine a world where, in addition to endorsing (or being inclined to endorse, or being at ease with) all of my desires, they are all fulfilled. Not only am I at liberty to climb any hill I wish, but I can travel there on free public transport, will receive good medical care if I fall off said hill and break my head, and so on. On this picture, I am not only deciding what kind of life I want to live, I am actually getting to live it - setting my goals and achieving them. If one of structuralism’s problems is that simply knowing what you want to do isn’t enough, then assuming that the desires haven’t been formed by oppressive socialisation or the like, it looks as if someone who knows what they want to do and actually does it must be autonomous.

It should be fairly obvious how the procedurally-independent desire-satisfaction (PIDS) account purports to satisfy the conditions of autonomy.

1. *Authority*. I am authoritative with regards to my self-direction insofar as I “decide for myself what is valuable” (Colburn 2010, p19), where this decision is not the result of malevolent or subversive influence.

2. *Power*. I am powerful insofar as I “live my life in accordance with that decision” (ibid); and I am autonomous insofar as I am authoritative and powerful.²⁶

²⁵ I have not included a ‘pure’ desire-satisfaction account, where one is autonomous just insofar as one’s desires are satisfied, for two reasons. The first is that (so far as I know) no-one has seriously proposed such an account. This is likely thanks to the second reason: such an account would imply that people acting under addiction or psychological compulsion, or with clearly adaptively formed preferences, would all be autonomous if they got what they desired. Such a theory on its face fails to capture either the power or the authority requirements given at the start of this paper, and so is of no interest to me.

²⁶ Colburn does not refer to his account as procedurally-independent desire-satisfaction (understandably), but I am confident that the name does no violence to his theory. Raz (1986) also holds a view not dissimilar to this, although his account is much friendlier to

Such a theory gives us a fairly straightforward account of what it is to be powerful and authoritative, and in addition seems to capture a great deal of what we mean when we talk of the nature and value of autonomy.

But just as we can critique structuralist accounts for allowing the “wrong kind” of desire harmonisation, such as when Friedman’s angel of the house suppresses her first-order desire to smash the dinner plates, desire-satisfaction or freedom of action accounts seem vulnerable to counting any kind of life as autonomously lived so long as the agent truly (madly, deeply) desires to so live it. While it is certainly difficult to construct a scenario whereby an agent could genuinely desire their own enslavement without there being oppressive socialisation or adaptive preference formation at work, it does not seem conceptually impossible that such a scenario could come about.²⁷

Not all autonomy theorists see this as a problem, of course: if it is in fact the case that someone has come to this desire free of influences which would render it inauthentic (implausible as that may be), such theorists would argue, then not only is fulfilling this desire compatible with their autonomy, it would be a limitation of their autonomy to deny them fulfilment. It seems, in fact, that there is no necessary reason why we might not be obliged to help the would-be voluntary slave to find a master.²⁸

With this in mind, it is helpful to consider why our intuitions (may) rebel against the notion that the voluntary slave is autonomous, not least because what seems to me to be the critical factor - the inability of the agent to thereafter fulfil any desires other than the ones which happen to coincide with the slaveholder’s wishes - suggests the third, social-relational, component required for a useful conception of autonomy.

perfectionism.

27 Although probably very unlikely sociologically, and I would be perfectly happy with a successful argument that ‘voluntary’ slavery is in fact an incoherent idea.

28 Colburn has indicated in discussion that he might be committed to this view, for example (though he asserts that there *will* almost invariably be good reasons not to find a master for the would-be slave). His 2011 paper ‘Autonomy and Adaptive Preferences’ also coheres with the view: so long as the desire for enslavement was the result of conscious character formation rather than covert influences upon our preferences, the voluntary slave is autonomous.

Imagine the agent who somehow comes to authentically desire their own enslavement, and imagine that this desire is satisfied. At some point afterwards, they realise just how limited their life now is; they lose the desire to be commanded by another, and develop a desire to be free - a desire which is inescapably frustrated by their enslaved status. The desire-satisfaction theorist, of course, can merely shrug and observe that they are not *now* autonomous, because they have authentic desires which cannot be fulfilled; but when they were content with their lot, when the life of the slave was one they desired, they were autonomous.

But there is a basic implausibility in this defence. By retreating to a fundamentally internalist standpoint - that is, committing itself to the claim that our autonomous status can change without any external change - it suggests that the slave becomes non-autonomous at precisely the point where they desire to be free, rather than at the point where their freedom is restricted.²⁹ In turn, this suggests that the best way to guarantee our autonomy is simply to aim low: the more easily-satisfied our desires, the more autonomous we are likely to be.

The obvious response, and one that we have already seen usefully deployed elsewhere, would be for the PIDS defender to show that “aiming low” would ultimately be a form of adaptive preference formation (and as such, something we already have the conceptual resources to explain without harm to the theory). This, however, will not work here: any such attempt is bound to define adaptive preference formation either so broadly as to include almost any desire, or so narrowly as to rule out paradigm cases of heteronomy.

If, for example, the defender of desire-satisfaction moves to suggest that aiming low would be an adaptively-formed preference because it would not survive the removal of the limitation (namely that thorough-going deep desires are likely to be frustrated), then we must ask why my preference for being fairly mobile is

²⁹ It also suggests a bizarrely dynamic flip-flopping account of autonomy: depending upon how my desires interact, I can be autonomous one day, non-autonomous the next, and then return to being autonomous the day after that, without there being *any* change in my physical circumstances or social standing (this observation is also made in Dworkin (1988), p16). This thought also serves as one motivation for the episodic/programmatic division to be examined later.

not adaptively formed.³⁰ After all, plausibly at least, I would be entirely indifferent to the difficulty of getting to a location by foot if I was not limited by gravity, inertia and the like. Given the ability to move from place to place by whatever means I liked, it strikes me that I might well opt for flight, or perhaps teleportation - but this cannot mean that my admittedly weak desire to be reasonably physically fit is adaptively-formed. If, in other words, the limitation of my being a (gravity-bound) human was removed, then perhaps my desire for fitness would vanish too. But this on its own seems an absurd reason to claim that my sporadic efforts at exercise are the result of an adaptively-formed preference.

There may of course be other reasons to desire fitness, but crucially there need not be - it seems entirely reasonable that someone could wish to be *fit* (rather than healthy) purely in order to move about more easily. If the brute fact that we are earthbound is not enough to render fitness desires adaptive, then how can the equally brute fact that many complicated or long-term desires are likely to be frustrated or at least difficult to fulfil mean that a higher-order preference for more easily-satisfied desires must be adaptively formed?³¹

If, conversely, we restrict the definition so that fitness desires are not counted as adaptive by dint of our physical constitution or limitation, then it is difficult to see how the desire for easily-fulfilled desires *is* adaptive. If we remove or at least loosen the restrictions on desire-fulfilment in our hypothetical world - so that all desires which do not contradict physical laws can be satisfied relatively easily - then while it might be a bit odd to desire easily-fulfilled preferences (odd because in this context the desire specifies nothing), it certainly isn't obviously incoherent.

In fact, given the new stipulations, the specificity of the organising desire makes it look very like a curiously self-effacing but perfectly conceivable desire that "I will desire only those things which are possible" - just as I might, if forced to

30 This account of adaptive preference formation is suggested by, e.g., Dimock (1997).

31 Come to that, such a picture of adaptive formation seems to suggest that if we have exclusive desires and a desire that they weren't exclusive ("I want to go to place x at time t", "I want to go to place Y at time t", and "I want it to be possible to go to both x and Y at time t", for example), the third must be adaptively formed because it wouldn't survive the removal of the law of non-contradiction. This seems odd.

seriously consider long-standing frivolous desires (for a starship, say) in the real world, conclude that there is no point in desiring what I literally cannot have.³² Indeed, perhaps the desire only seems curious in the context of our hypothetical world of maximal desire-satisfaction, because it seems psychologically odd to wish only for what you can get when you can get more or less anything.

That is to say: if *all* desires that can be fulfilled in the hypothetical world are fairly easily-fulfilled, and this means a preference for easily-fulfilled desires just equates to a desire not to have impossible desires, then it seems entirely possible and indeed sensible for this preference to withstand the removal of limitations which brought it about (namely that some desires are far more likely to be frustrated than others). As such, it need not be adaptively formed, and the “aiming low” criticism cannot be defused by reference to the formation of the desire.

The objection will be raised here that there is a key difference between my desire for a starship and someone else’s equally difficult to fulfil desire for (say) control over the direction of their life, namely that my desire is frustrated by factors which cannot be defeated through correcting injustice - given the various historical, technological and medical conditions which *actually* obtain, it is simply not going to happen.

The agent’s desire for self-direction, however, is not like this - given the various historical, social and economic conditions which actually obtain, it is still possible (not to mention morally obligatory) for their desire to be fulfilled.³³ To suggest that my science-fiction desires and their desires to be free are equally fantastical seems not only mistaken, then, but deeply objectionable. The comparison may be thought to invite or imply the conclusion that we are equally

32 Although the limitation is contingent and it makes sense to speak of me desiring a starship insofar as this desire incorporates the desire for other requirements which would make the desire possible, I could not seriously desire it in the same way that I could desire a cheese sandwich, a beer, or a trip to Dublin - given contemporary levels of technology and the average human lifespan, then barring a technological singularity (by definition impossible to predict), it is simply impossible that my desire for a starship will be fulfilled.

33 This is not to suggest that the *only* important difference between the desires is one of feasibility; obviously a deep desire for self-direction is far wider-ranging and has more normative weight than a vague wish that reality was more like Star Trek, but these differences are not yet relevant.

as culpable in our own frustrated desires as those who frustrate them; “well, if you *will* go around wanting to be free...”.

But this is *precisely* the problem that the comparison is intended to highlight. Clearly it is vastly more unfair and unjust that the agent is oppressed than that I have no starship, but merely casting it in terms of frustrated desires fails to fully capture how the loss of autonomy occurs. While it is the case that the structural barriers to my starship are of a generally different kind than the structural barriers to the agent’s freedom, they are still structural; the agent may be (and usually *is*, at least individually) just as powerless to change the political structure which oppresses them as I am to change the technological structures which so consistently fail to crack faster-than-light travel.

This being so, and given what was said about adaptive preferences earlier, the desire-satisfaction theorist seems committed to arguing that I am rendered less autonomous by my lack of starship, and that the other agent would be autonomous if it wasn’t for the unfortunate fact of their desires. The first of these claims seems suspicious at best (although much of this suspicion is probably sociological rather than philosophical, a point which will become more important later); the second, outright oppressive.

It is worth re-stating that the requirement for non-adaptively formed preferences probably catches most if not all actual (or closest-possible-world) cases where folk come to drop desires for freedom, self-determination and the like. My suggestion is not that, for example, a desire-satisfaction theorist holding this condition is always unable to explain why the angel of the house is not autonomous, but rather that there are tweaks - and not absurd, fantastical tweaks, either - to such cases which will block these kind of explanations; and that this in turn suggests that the mechanism which affects autonomy is not authenticity of desires but some other external operator.

So far, neither internalist accounts nor theories concerned with the fulfilment of authentic desires have given a satisfying account of autonomy’s structure. Desire-satisfaction accounts either deny the loss of autonomy (if the prisoner is voluntarily incarcerated, for example) and thereby fail to meet the power criterion, or are committed to suggesting that it is contingent on the agent’s

desires. The foregoing argument has shown that the latter option is unworkable, leading as it does to the counter-intuitive conclusion that the best way to increase autonomy is to desire as little as possible.

Procedurally-independent desire-satisfaction accounts, the most plausible form of desire-satisfaction theories so far, seem at first to meet the power and authority conditions. Indeed, in some sense they *do* meet them: we're authoritative over our desires and have the power to pursue them. In fulfilling the conditions however, the PIDS account reveals that we're still not being external enough about them. It seems to me, for example, that Colburn is committed to the flip-flopping autonomy (where I could cycle rapidly between autonomy and heteronomy without any change in my external circumstances or social positioning) which Dworkin and I have criticised as implausible, as well as to the autonomy of the voluntary slave.³⁴ The weak authority condition I mentioned earlier should be replaced with a stronger one: it is not enough to be authoritative over one's desires. Similarly, the conception of being powerful over one's life needs sharpening, for worries about domination lurk just off-stage. Being able to pursue one's desires does not seem to meet the power criterion for autonomy if, as in Pettit's (1996) famous example, we are only able to pursue our desires in virtue of somebody else's refraining from interference with our lives.

In directly addressing the first prong of the desire-satisfaction response, the claim that a voluntary slave who remains happy with their lot *is* autonomous, I also set out the third option for giving an account of autonomy's constitution: the relational (or, following Oshana, social-relational) theory of autonomy. Latterly, I examine the domination worry in more detail, arguing that the relational account can explain precisely why, 1) the weak authority condition is

³⁴ It has been objected that there is, in fact, an external change here: namely, that your desires are no longer being satisfied (we might cast this as something like a "desire-frustration relation" obtaining where no such relation obtained before). I'm not wholly convinced that this *is* a meaningful external change, but I can accommodate the charge by instead arguing that the external change (if there is one) is *entirely* dependent on the psychological change, and can happen without any other external changes in terms of how powerful and authoritative we are. The social relations that we stand in remain the same, and we have just as much power over the direction of our lives as we did before; all that's changed is that now we're not happy about it.

insufficient, and 2) The satisfaction of procedurally independently formed desires is consistent with being dominated (and thus non-autonomous).

3.2 Relational autonomy

The most plausible components of the previous accounts both suggest that at the very least, we need power over ourselves and the direction of our lives in order to be autonomous. Although the power does not need to be the improbable total self-mastery of the Randian hero, it does need to extend far enough that we can recognise and reject inauthentic desires or oppressive socialisation, and that we can both pursue *and change* authentic desires. Unless we have a fairly explicit desire to be regarded as an independent source of values then the loss of power cannot inhere in the frustration of desires as such.³⁵ To return again to the voluntary slave example, they are not empowered by the satisfaction of their slavery desire, nor would they be disempowered by its frustration. If the desires play any role here, it is only insofar as their fulfilment or frustration affects our social relations. This is a reversal of the more usual (particularly in proceduralist accounts) claim that our relations with others only impact our autonomy insofar as they adversely affect our ability to form authentic desires or to carry these desires out. In the rest of the chapter, I show that taking social relations seriously as constituents of autonomy is implied by the particular failures of non-relational accounts, and make the positive case for the relational view.

To re-state the social-relational claim more positively: our relations are important not only because they can causally influence our autonomy (as in when some set of relations allows me to or prevents me from fulfilling a desire), but because they partly *constitute* our autonomy: a large chunk of what it is to be autonomous is to stand in particular relations or kinds of relations. The voluntary slave is thus non-autonomous because, regardless of how heartily they endorse the values which control and shape their lives, their endorsement is not the deciding factor in what values play this role. The relations that the slave stands in are those of domination and subordination to others, and it is purely contingent that their values and the values of their dominator are congruent; or,

³⁵ We might, of course, have such a desire, but equally well we may not - indeed, we might be motivated (perhaps out of a belief that the ultimate source of value is god, for example) to reject such a desire even if we did have it.

to borrow Oshana's neat precis, the direction of the voluntary slave's life "coincides with what she desires, but is not dependent on what she decides" (2006, p61).

The relational interpretation of the voluntary slave case avoids the flip-flopping of autonomy which bedevils desire-satisfaction accounts. If the agent endorses the enslavement one day and repudiates it the next, their happiness will clearly be affected, but without a concomitant change in the relations they stand in, their autonomous status remains unchanged. The relational conception also offers us a simple explanation of why they lack power and thus autonomy: the voluntary slave may very well be living her life in accordance with her decisions about what is valuable, but she lives so *only* on the whim of the slave-owner, for the relations that she stands in are not those which grant her robust powers to choose - and change - the direction of her life. Similarly, the powers that she does have are vulnerable to being altered or suppressed should the dominators so wish it.

Imagine that the reasons such an agent has to opt for enslavement are based in some religious concerns; "only religious authorities of such-and-such a composition are proper judges of how to direct a person's life, because what is valuable in life is service to god(s) and these authorities are best placed to know the will of god(s)". This reason lasts right up until the authorities in question undergo some dogmatic schism which results in the slave being controlled by a group she strongly disagrees with; she can no longer be said to be living her life in accordance with what is valuable to her, but her actual powers have changed not at all.

Perhaps, though, this case trades too strongly on our sociological suspicion about whether the agent here is *really* procedurally-independent, or whether such a course of action is rational according to the standards and norms that govern self-direction. Here is a different example. The agent now regards survival, rather than religious service, as the primary or sole object of value: she will favour all and only actions which give her the best chance of seeing the next sunrise.³⁶ This surely cannot be irrational, particularly (though not only) if we

36 Clearly, the agent is a thorough-going Hobbesian.

specify that disaster conditions obtain. To this end, she submits to enslavement along the same lines as the other cases we have posited; and again, she seems to be autonomous on the PIDS account until the slave-owners begin to act in ways contrary to her survival. But, again, she is no less powerful than she was before the slave-owners began to interfere; the problem is not that somebody has interfered with her self-direction, but that she is entirely powerless to stop them.

In contrast, an agent who does stand in relations that make her powerful and authoritative is not prone to arbitrary interference in her life; should she, for example, come to believe in a different religion or interpretation of her religion then there is nothing which can force her to continue a life of service to the rejected church. Indeed, if she stands in the relevant relations, then *only* her faith or belief can compel her to continue a life of religious service in the first place.

The shape of the forthcoming argument regarding relational standing and power over one's life should now be clear to anyone familiar with Philip Pettit's work on republican freedom, but for clarity and comprehensiveness' sake it is worth rehearsing Pettit's own argument about domination and power before turning it to my purposes.

3.3 Domination, Power, and Why the Weak Authority Condition is Too Weak

Pettit (1996, 1997) presents the hypothetical case of a slave who has a good working knowledge of the psychological quirks and tics of the slaveholder, such that they can (with a fairly high degree of success) effectively pursue their own desires and evade much or all attempted interference with their lives. To give a concrete example: they know that the slaveholder is very fond of port and cheese on a Friday afternoon, and that if they're given it in sufficient quantities, they will pass the rest of the day in satiated quietude. Thus, if the slave can contrive to make sure that the slaveholder is stuffed with food and drink, they have most of Friday to themselves - there will be, albeit temporarily, no interference with their lives.

As Pettit remarks, it would seem very odd to claim that freedom from interference made the slave meaningfully free. As it happens, they won't be interfered with, true, but if the slaveholder changes his habits then the slave's relative Friday freedom is lost and they have no recourse against it. What is necessary for freedom, Pettit argues (1997, pp64-68), is that we are free from *domination*: nobody should be in a position such that they can arbitrarily interfere with our lives without our having any recourse against them. In the example above, the slave has no recourse against the slaveholder's suddenly deciding to interfere with their Friday activities. To borrow Oshana's phrasing again, their lives have (in this respect) coincided with their desires, but have not been dependent on their desires (op. cit., p61). Rather, they have been dependent on the whims of the slaveholder, and so when the whims of the slaveholder stop coinciding with the slaves' self-direction, the slaves can no longer self-direct.

It is easy enough to draw the parallels between freedom in Pettit's account and autonomy here: we are autonomous insofar as we are powerful and authoritative, and neither power nor authority are consistent with being dominated in Pettit's sense. To stand in autonomy-constitutive relations of power and authority is to stand in relations which are hostile to arbitrary interference with one's life.

What, though, is unique or important about the relational view here? Why can't the PIDS theorist, for example, steal the relational account's thunder by embracing non-domination as being required for procedurally-independent desire-satisfaction?

Let's take power first. Can the slave satisfy desires that she has formed under conditions of procedural independence? It seems obvious that she can, *dependent on what those desires are*. Even if her desires are not of the self-effacing kind that ought to make us worry about the independence of their formation, she can clearly fulfil at least some of them if, like Pettit's agent, she is subject to a slaveholder with no interest in frustrating those desires. But this surely does not motivate us to say that she would be autonomous, and for precisely the same reasons that Pettit's slave is not free. Note that if the slaveholder's whims change, such that the enslaved agent can no longer pursue

those desires, there has been *no relational change* - she still stands in dominated relations whether or not there is active interference. Consistency demands that either she is autonomous in both cases or in neither, and surely nobody would assert the autonomy of somebody whose desires were consistently frustrated by the arbitrary whims of another. If it is not the as-it-happens frustration of the desires, but the structural inability of the agent to pursue any desires that don't line up with the slaveholder's, then what makes them non-autonomous is their social-relational standing. To put it as plainly as possible: if we agree that domination is incompatible with power over one's life, then (since desire-satisfaction is obviously compatible with domination) we cannot think that power over one's life consists simply in satisfying desires.

What about authority? Recall that I have already made vaguely foreboding noises about the ability of the *Weak Authority Condition* (roughly, "I have some minimum level of self-respect, and regard many or most of my actions as being explicable by this respect for my agency") to do the job required. One obvious problem for the WAC would be radical misunderstandings of one's own actions or psychology. Imagine a Manchurian Candidate, where the brainwashed agent believes that assassinating the President is explained by her desire to kill the President and her belief that she is an appropriate source of values. She's just *wrong*; the desire has been implanted in her in such a fashion that she thinks it's explicable by reference to her values, but of course this belief is itself the result of psychological manipulation.

Now, adaptive preference formation concerns will obviously disbar this as an example of autonomy - and rightly so - but exactly the same kind of problem applies when the agent is wrong not about why *she* acts, but about how others understand her actions. Take the following example: an enslaved agent is told to instruct the slaveholder's children in some or other subject. As it happens, this is something that she would want to do anyway, and if she was asked why she set up a lesson in such-and-such manner, or chose this or that subject, her reply would involve her own values. But so far as anybody else is concerned, her actions are explicable just by reference to the demands of the slaveholder; and, importantly, if there is a clash between her and the slaveholder on some

pedagogical point, any claim on her part that involved her own values would be regarded as irrelevant.

Another, perhaps less contrived, example may illustrate the point further. Take a citizen who wants to be a teacher (for reasons of their own), and who is supported in this goal. They have - by hypothesis - the power to direct their life; they certainly seem to have a minimal level of self-respect and the belief that becoming a teacher is explicable by reference to these values. But now add that they are supported in this goal only because, and only insofar as, they will aid the state in some aim or goal which has nothing to do with their own values. Perhaps, for example, their support is dependent upon keeping their teaching within certain ideological boundaries, or upon churning out so many officer cadets per year. Again we see that although the agent regards their being a teacher as being explicable by reference to their principles, such an explanation would gain no traction in a dispute with the governing body. That is to say, where they see the question "why did you teach this class in such-and-such a manner?" as being at least partly answered by "my belief that this manner of teaching furthers a deeply-held principle of mine", their interlocutor will not. If their teaching career is solely dependent upon their efficient replication of some government ideology, it seems obvious that they are *not* being supported in virtue of their agency; there is no externally recognised connection between their values and their actions.

The lay of the land is as follows. If we think that autonomy is valuable because, ultimately, agency is valuable (and agency is valuable because it allows us to engage in conscious and creative manipulation of our environment in accordance with our principles and goals), then it seems very odd to say that autonomy can consist in a life where the prime explanatory factor is some other body or institution's goals. Something crucial about authority is missed by such a life; just as the canny slave's desire-satisfaction might be consistent with domination rather than indicating freedom, somebody's being allowed to perform or supported in performing some act (or series of acts, etc.) is not in itself enough to show that the support has anything to do with respect for them as an agent.

To this end, I suggest that the WAC must be replaced with the imaginatively-titled Strong Authority Condition: in order for an agent to be authoritative over their life, they must, 1) Have some minimum level of self-respect, 2) Regard many or most of their actions as being explicable by reference to their goals or values, 3) *Be generally regarded by others as an appropriate source of values.* This third sub-condition, as is immediately obvious, is not something that can be conceptualised in terms of satisfied desires, but rather consists in having a certain social recognition or status - in other words, standing in particular social relations. The question of what kind of relations these might be is a substantive one, and as such best left till the next chapter; but formulating the Strong Authority Condition (called SAC from here on out) in this way explains exactly what's gone wrong in the "enslaved teacher" and "government puppet teacher" cases. Although they see their self-direction as being explicable by reference to their goals, nobody else does; and it cannot be the case that we recognise an agent as an appropriate source of values - that is, recognise them as authoritative - if we do not think that their actions should be construed in terms of their desires and values.

My account, in summary, gives the following criteria for being powerful and authoritative over one's life:

1. *Power.* We are powerful over our lives if we stand in sets of social relations such that we can effectively pursue our goals - although these relations must be consistent with maximal *equal* autonomy; standing in relations of "being a tyrant" is, for reasons which will become apparent next chapter, inconsistent with autonomy.
2. *Authority.* We are authoritative over our lives if we satisfy the strong authority criterion: we have a minimal level of self-respect, regard many or most of our actions as being explicable mainly by reference to our goals and values, and are generally regarded as an appropriate source of values. This third sub-condition is external and relational, while the first two are primarily or entirely psychological.

This account explains why Sinead is not autonomous: she is relatively powerless in addition to not being regarded as an appropriate source of values. It captures

Seamus' heteronomy by noting both his inability to self-direct, and the obvious refusal of others to acknowledge his authority over his life. Finally, my account is in a position to show why domination *is* inconsistent with autonomy in a fashion that, I have argued, PIDS is conceptually incapable of doing.

There is, however, a significant ditch left to jump. Specifically, the account as I have set it up seems prone to skipping over self-direction entirely; so long as there's nothing unjust at work, it appears, whether or not I actually can pursue (or am pursuing) my values is somewhat orthogonal to the question of whether I'm autonomous.

3.4 Worries About Voluntary Slavery and "Blamelessly" Frustrated Self-Direction

There are two kinds of self-direction worry for the relational account addressed in this subsection. The first is the familiar problem of whether the voluntary slave is autonomous, and the second concerns self-direction frustrated without any apparent injustice. Both of these worries are related - in both cases we are confronted with an individual who is apparently autonomous and yet cannot self-direct in some or all realms - but they are distinct. The case of the voluntary slave illuminates important distinctions between autonomous actions and an autonomous life, while the problem of frustrated self-direction without injustice stresses the relational account's focus on social and political structures rather than desire satisfaction. In both cases, I argue that the relational view gives the right answer: the voluntary slave is not autonomous because the relations they stand in do not constitute their being powerful and authoritative, while the agent whose self-direction is (blamelessly) frustrated still stands in relations of power and authority over the course of their life.

To take them in order, recall the hypothetical example of the voluntary slave: an agent who has some (unspecified, but let us presume authentic) desire which requires them to give up control over their life to some other(s), and who successfully pursues this desire. On the view I've suggested, they cannot be autonomous as slaves - their realm of self-direction is delimited by another and they seem to be fairly comprehensively dominated - but then this is exactly how they want their life to go; their status as a slave is one that they genuinely endorse.

Now, I am not interested in here defending the claim that they are non-autonomous (the previous subsection should anyway make it clear why I take the claim to be true), but a case wherein a previously-autonomous individual satisfies an authentic desire and thus becomes heteronomous does invite the question of how we should think of the act which leads to abrogated autonomy. When an agent enters into voluntary slavery, do they autonomously commit themselves to heteronomy, or does the nature of the act somehow render its performance heteronomous?

The latter option seems blocked for relational theorists, given that carrying out the act under the stated conditions *requires* that the agent stand in relations of autonomy - if they weren't able to self-direct, they wouldn't be able to direct themselves into slavery. And given the external emphasis of relational accounts, an attempt to disbar the agent or the act from autonomy because of (non-adaptive) desires they have would surely be inconsistent. That is to say: claiming that the agent will not be autonomous when they are a slave, because they lack the relational standing crucial to effective self-direction, means that we cannot simultaneously argue that they are non-autonomous when they *do* stand in such relations but wish not to.

This, then, appears to leave open only the path of arguing that the agent is indeed autonomous at the point where they volunteer for slavery, in which case we might well be tempted to ask again what the relational view is bringing to the table. If the account says that the agent is autonomous because being a slave meets their deep values, then it just looks like desire-satisfaction with some added bells and whistles. If, on the other hand, the agent is not autonomous because they're performing the wrong kind of self-direction, then the objection will be that we're being arbitrary about when the power and authority to self-direct constitutes autonomy.

In responding to this objection, the distinction between autonomous choices and autonomous lives, or between 'episodic' and 'programmatic' autonomy, becomes crucial.³⁷ Episodic autonomy is most easily thought of as being

³⁷ I use Meyers' (1989) "episodic/programmatic" nomenclature for this distinction rather than "local/global", but I think the terms are equivalent. Similarly, both Raz's notion of pervasive autonomy, and his claim that "the autonomous person chooses his own profession or trade. He

instantiated by decisions about what kind of fizzy juice to buy, or which route home to take; if the shop is out of Irn-Bru, or if roadworks block my preferred path, some desire of mine has been frustrated. However, it is not at all clear that my social relations have changed. I still stand in the kind of relations that allow me to choose *x* or *y* drink or path, but on this particular occasion *x* (or *y*) are not options on the table. It is also notable that an agent could be episodically autonomous under conditions of extreme oppression - think of a Guantanamo Bay detainee who chooses to walk around the exercise yard clockwise rather than anti-clockwise, or a person of colour in 1950s England deciding between two kinds of beer. We would patently resist attempts to suggest that these agents stand in the kind of relations which characterise the maximally autonomous person - you cannot pursue life to the fullest in an internment camp, nor while at constant risk of systemic, verbal and physical abuse - but there is obviously a sense in which they *are* in control of how their life goes, albeit a limited sense. Here, then, we can speak of the agent as episodically autonomous but programmatically non (or minimally) autonomous: free to choose this direction rather than that, or this pub over that one, but not free to be safe in either case without conforming to norms that are hostile to their having maximal equal autonomy.³⁸

This is, admittedly, a somewhat simplified picture. There can be instances of episodic autonomy that have profound effects on the direction of our lives: whether I decide to go to one job interview or another, for example. Here, the exercise of my episodic autonomy has obvious and far-reaching consequences for my programmatic autonomy; nonetheless, it still seems to be the case that programmatic autonomy is more than a mere “adding up” of instances of episodic autonomy. Although episodic autonomy *can* be important for our programmatic autonomy, it will not always be so.

may be denied the chance to cut down trees in the next field without any diminution to his [pervasive] autonomy” cohere neatly with the view of programmatic autonomy I put forward. See Chapters 4 and 6, in addition to Meyers (op. cit.), Oshana (op. cit.), and Raz (1986, p409).

38 I explore the question of why interfering with an agent’s self-direction (even if that direction would express their deeply-held principles) is not simply equivalent to limiting their autonomy in the next chapter.

As the foregoing paragraphs suggest, programmatic autonomy (by partial contrast) involves standing in the relations necessary to pursue those desires and goals which are central to our self-conception. To restrict programmatic autonomy is to non-trivially interfere with the direction of someone's life and, indeed, with the constitution of their social self - denying an agent freedom of self-expression is in some sense to try and change who that agent *is*.³⁹

This accounts for the intuition that to claim I am non-autonomous because I am not the captain of the *Enterprise-E* is frivolous. Unlike the slave, I have substantial power to direct my own life and pursue or change desires central to my self-conception - where I do not (for example, with regards to where I live and travel, the relevant relations I stand in are fairly limiting, albeit far less so in comparison to many others), it seems entirely correct to say that I am consequently lacking in autonomy.⁴⁰ Thus, fairly straightforwardly, the voluntary slave is (or may be) autonomous when she decides she wants to be a slave, and is not autonomous when she stands in enslaving relations; and the change is explained very simply by reference to her actual power and authority rather than how she feels about her power and authority.

Still, something about the problem of my thwarted desire to go star trekking *does* stick, and deserves attention at the structural rather than substantive level: the worry about blamelessly frustrated self-direction. We can think of this as a kind of mirror of the voluntary slave example. Instead of worrying that we might get a false negative by claiming heteronomy of an agent's authentic self-direction, the problem here is one of a potential false positive, where we ascribe autonomy to somebody whose self-direction appears crucially limited

39 I am inclined to think that there's an asymmetry here with episodic autonomy: episodic autonomy can be (though it often isn't) vital to programmatic autonomy and "leading a life of one's own", while restricting the exercise of programmatic autonomy will invariably interfere with one's power to lead such a life. This isn't to say that the programmatically autonomous agent must be a go-getting individual - she may perfectly well lead a "timid, limited, and mundane" life, to borrow a characterisation from Oshana - but simply that restricting programmatic autonomy involves a change of relations which is not mirrored in any particular instance of limited episodic autonomy (unless that episodic autonomy is limited as a result of limited programmatic autonomy). If I am persistently unable to choose my preferred brand of fizzy juice, I still seem to be in control of my life; if I am persistently unable to choose which job to do, I am so much the less powerful over the direction of that life (even if I don't, as it happens, *care*).

40 Again, the "Douglass Objection" that the overcoming of obstacles might itself be thought to make one *more* autonomous will be addressed in the following chapter.

despite the absence of any malevolent interference - hence, “blamelessly” frustrated self-direction.

A familiar example might help to illustrate the point. Imagine again that I wish to be a teacher, that this role plays a significant part in my self-conception; and, further, that so wishing is taken as being reason enough for me to be supported through training, accommodation, grants and the like - in short, that I stand in all the appropriate relations to have power and authority over my life. Now, imagine that my self-direction is frustrated; not by sudden collapses in instrumental reasoning, or through malign interference, but because I’m just not very good. Perhaps it turns out that standing in front of a class makes me panic, or that I take student feedback personally, or whatever the case may be. In any event, it is clear to all concerned that my teaching career is a non-starter: I will not achieve that goal.

It certainly doesn’t seem that I have failed to meet any of the (plausible) internalist criteria: I am pursuing authentic desires, and no-one has hindered my ability to self-reflect. Similarly, the relations I stand in are largely the same as if I were a “natural” - my assessor is supportive, I have not faced unjust barriers, and I do not risk poverty or death in the event of failure. But still and all, there is something exceedingly odd about the thought that I might be autonomous and at the same time be *unable* to fulfil my self-conception.

I have placed the worry here, rather than as a substantive challenge in the next chapter, because it suggests at first glance that my account of autonomy as power and authority relationally conceived is missing some key structural feature. If this agent really is both powerful and authoritative, the critic will suggest, then either my conception of power and authority is incomplete or incorrect, or there is some third facet to autonomy. Just as Seamus and Sinead’s cases act as defeaters to previously discussed theories, the thought here is that an inability to account for the failed teacher is enough to suggest that the relational approach misses something absolutely key to autonomy. In other words, Sinead defeats structuralism; Seamus highlights the inadequacy of procedural independence; the voluntary slave hamstring PIDS; I must therefore explain why the failed teacher does not sink my relational account. Showing *that*

my argument gets it right in this case serves that purpose, while showing *how* it gets it right (hopefully) illuminates the argument further.

In pursuing these tasks, it will be helpful to recall why we are concerned with autonomy in the first place. We value self-direction because of the value of agency - we should seek to support others in their self-direction in virtue of having *prima facie* reasons to think that people self-directing is a good thing. The best conceptual method of evincing and respecting the value of self-direction, I have argued, is to think of the autonomous agent as one who relationally situated such that they are powerful and authoritative over their life.

This claim certainly seems in tension with arguing that the unsuccessful teacher is autonomous but it is not clear that this tension is anything but superficial (or perhaps another result of the autarchy/autonomy conflation). Consider that, in the failed teacher case, my wanting to be a teacher was taken as reason enough for me to train as a teacher, and that there were - by stipulation - no oppressive relations which prevented me from pursuing that plan of life. In other words, I was able to pursue something I valued deeply, because I valued it deeply; this sounds very like the kind of thing that an autonomous agent does. Again, Raz's remark that the autonomous lumberjack might not be free to work in this field rather than that is pertinent here. We cannot require that the autonomous agent is one who is successful in all she pursues. People fail to achieve their goals for numerous reasons, and some of those reasons seem to be inextricably woven into the tapestry (or trodden into the manky rug) of human life.

If a five-footer fails to achieve their goal of becoming a world-famous basketballer, we can commiserate with them and share their disappointment, but we do not think that they have thereby become less able to pursue values in general, or that they have been unjustly prevented from pursuing those values. There is, in short, a distinction that must be drawn between "rotten luck" and "rotten society", and having rotten luck is not in general enough to mean that one is non-autonomous. The worry that this is a change in autonomy without an external change - a possibility I have previously rejected - can be addressed in one of two ways. Either the relations I stand in *have* changed from, say, "would-be teacher" to "not a teacher" (think of a teaching student as opposed to a

retired teacher, or indeed an apiarist or shopkeeper) or the relations haven't changed *but neither has my autonomous status*. That is, if I really stood in "not a teacher" relations all along, unbeknownst to me, then my self-direction would always have been frustrated. Whether or not I am autonomous in that regard, I am no more or less so than I was a year ago.

Does this mean that we should accept the apparent oddity of an account which gives us, in the failed teacher, an autonomous yet persistently frustrated agent? I believe so - the pain is not so very great as to make the bullet-biting more difficult than is usual in philosophy. Autonomy as we have conceived of it is primarily a social matter: a question of how we ought to organise society given our purported respect for the value of agency. There is nothing that society can, or perhaps ought, to do about my being a hopeless teacher or too short to play basketball - except to ensure that I have opportunities to pursue other values, or those same values in other ways.⁴¹

Although I am increasingly convinced that something like this is the right response, its counter-intuitiveness notwithstanding, there is another response which grants the agent's heteronomy while still supporting the claim that such agents are not a problem for my theory.

The first move would be to admit the conceptual possibility of the failed teacher, but deny that the case is sociologically probable. The agent could, after all, work in other educational roles or find other ways to self-direct according to the deeply-held values which motivated her to pursue teaching in the first place. Recall again Raz's analogy of the lumberjack, whose autonomy is a matter of being able to lumberjack rather than cut down a particular tree. Imagine that they pursue the profession because they value being outdoors, doing physically

41 One might be tempted to take this case as supporting the idea that responsibility is the key to autonomy - again, see Colburn (*op. cit.*), but the notion can be crudely summarised with the claim that we're autonomous insofar as we can appropriately be described as the author of our actions (even if those actions are, say, an unsuccessful pursuit of some goal). This is not the place for a full rebuttal, but given the obvious relevance to the problem of frustrated self-direction it is worth saying something. Briefly, it seems to me that one's standing in relations which constitute power and authority over one's life is likely to make one responsible for (or constitute one's being responsible for) that life - the responsibility is a feature of, not a prerequisite for, autonomy. The relational account can thus encompass at least some responsibility accounts, whilst explaining that what makes the failed teacher autonomous is her relational standing rather than her responsibility for the unsuccessful attempt.

taxing work, and so on. If it turns out that they (for example) have a mortal phobia of the sound of falling trees, or an allergy to sap, and as such cannot be a lumberjack, then they are still not prevented from doing a job that will allow them to be outdoors, etc.

This may appear to be defining the problem out of existence by way of ever more abstract conceptions of what it is to pursue one's values, but that is not the intent. Instead, observe that it would seem very odd for somebody's continuing to have as a central part of their self-conception something that they know they cannot do. In certain cases, where this inability is the result of structural malevolence, it would be significantly less odd - "I should be a teacher, and I would be if it wasn't for the institutional racism" is a complaint with a genuine force that "I should be a teacher, and I would be if I wasn't crap at it" patently lacks.⁴²

All of this suggests a distinction between self-direction frustrated by malign influence and self-direction frustrated by 'faultless facts' about the world (I'm too short to be a basketball player, too shy to be a teacher, too pre-Singularity to fly a starship) which is relevant here. It is one that can be tentatively identified through considering how we would respond to "I would be an x if it wasn't for y" type statements - if y is some faultless fact then ultimately we have a kind of rationality failure with regards to the agent's self-direction. Either the failed teacher is not a worry for my account because folk don't, as it happens, tend to centre their self-conceptions around values or occupations made impossible by faultless facts; or it is not a worry because such an agent's continuing to do so would make them heteronomous (or less autonomous) by dint of a problem with their instrumental rationality. It would be, on some level, their failing to be authoritative over the direction of their lives, analogously to how my authority over what film I want to watch would be undermined if I selected *Wuthering Heights* on the basis of an affinity for comedy.

One obvious objection is that we are highly likely to have disputes over whether something is a faultless fact or not - so be it. There would certainly have to be

⁴² Similarly, "I should be a starship captain, and I would be if it wasn't for our technological level and certain universal constants" is obviously not a complaint we should take seriously.

some criteria to that effect, but that is a substantive task for elsewhere. I can, however, offer a tentative sketch of what these criteria might be. If the feature or fact which prevents us from pursuing some goal does not involve arbitrary interference (or standing in relations which make us vulnerable to such interference), then it is likely to be what I've called a faultless fact.⁴³ To give a pair of contrast cases, my being unable to be a starship captain does not entail that I stand in dominating relations - it is, as suggested, entirely consistent with my being powerful and authoritative over the direction of my life - but someone's being prevented from pursuing a teaching career because of racist biases or hiring practices is (one instance of) what it is to stand in dominating relations. Explaining why I can't command the *Enterprise* involves no reference to unjust and arbitrary relations of power; explaining why the discriminated-against agent can't be a teacher does.

With these criteria in hand, denying that *in general* there can ever be times when an autonomous agent might be frustrated or unfulfilled would be - to emphasise the point - to give an account of autonomy as autarchy, and hence cease to address the functional role of autonomy as a concept explaining the value and practice of self-direction. We want, to restate the goal, a story of why and when self-direction is valuable; merely asserting the primacy of autarchy does not tell that story.

I have suggested, in response to the worry that my account still seems to suffer from autonomy being blocked or limited by purely internalist factors, two possible options.

We might, first, concede that since autonomy as a concept is concerned with human activities rather than the activities of perfectly rational and near-omnipotent agents, the autonomous agent can still find her self-conception butting up against an unfortunately but blamelessly unsympathetic world. The failed teacher is only heteronomous if she finds herself disempowered or treated with less agential respect as a result of her failure; it is not inconsistent with autonomy to be unhappy and at a loss with what to do with one's life.

⁴³ This standard for distinguishing between merely unfortunate faultless facts and instances of injustice was suggested by Marina Oshana.

Alternatively, we could say that the failed teacher is heteronomous, but is heteronomous for the same kinds of internalist reasons that we have already baked into the account. If I were to suffer a brain haemorrhage and fall into a coma, my desires would be frustrated without external interference, but this seems unproblematic.

Both of these responses follow similar lines, but can be thought of as offering two variations on why the failed teacher is not a problem for the relational account of autonomy: either they are not heteronomous, or their heteronomy is not an instance or result of the sort of thing that should worry the relational theorist.

Conclusion

If, as I have previously argued, the autonomous agent should be thought of as one who is powerful and authoritative over the direction of her life, then internalist accounts - theories which locate autonomy in the relationship between psychological states - plainly will not do the job. Seamus the prisoner can identify with his desire to go hill-walking as hard as he likes, but this will not make him autonomous until that desire gives him a normative claim (in other words, until he is authoritative) and until he actually has the power to do so.

Some measure of higher-order identification seems necessary for *authenticity*, although the Sinead case shows that this identification must itself meet an external condition (that of procedural or historical independence), but autonomy as a concept that illuminates the value of self-direction is clearly not served by accounts which limit themselves entirely or primarily to psychological orderliness.

Procedurally-independent desire-satisfaction can adequately explain both the Seamus and Sinead cases: Sinead is not autonomous because she is not procedurally independent, and Seamus is not autonomous because his desires are unsatisfied. Although an improvement over internalist accounts, PIDS fails to properly heed the power requirements of autonomy: we can become less autonomous, without in any way becoming less able to direct our lives, just in

virtue of new desires. Further, PIDS allows for the autonomy of clearly dominated individuals: in addition to failing the power requirement, this also makes the agent's authority impossible, for how could an authoritative agent be denied even *prima facie* status as an independent source of values?

Only the relational account gets the right answer in all these cases. If we value autonomy because we value agency and self-direction, then power must be understood not as the mere absence of interference, or ability to evade it in pursuit of goals, but rather as being situated within relations and material conditions which tend towards maximising any agent's ability to self-direct according to their values. Similarly, authority cannot simply be the belief that one's actions are explicable by reference to one's values; being *recognised* as an independent source of values is a necessary criterion for being autonomous within society.

I have argued that, counter-intuitive as it may at first seem, a certain amount of frustrated self-direction is perfectly acceptable for this account: structuring society to give everybody the best possible chance at autonomy cannot require making sure that everybody gets what they want (even if what they want is entirely unobjectionable) all the time. If we are concerned with autonomy, our first worry should be oppression, not hard luck; and the relational account takes these worries on board in a way that none of its competitors do.

To be autonomous, then, is to be powerful and authoritative, where the external criteria for power and authority are constituted relationally; in the next chapter, I address the substantive task of laying out and defending the key relational constituents.

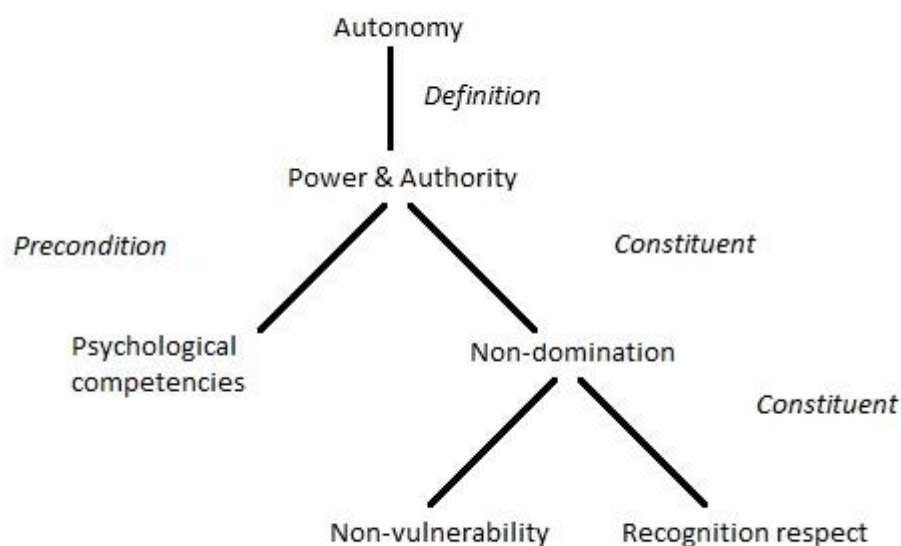
4.

Autonomous Relations, the Personal and the Political

Introduction

In the preceding chapters I have argued that personal-political autonomy is not best understood as self-mastery or unrestricted choosing. Instead, we should conceive of it as constituted to a large degree by external factors: the social and political relations we stand in. Only by understanding it in this way can we defend it against charges of hyper-individualism and make sense of paradigm problem cases for autonomy theorists.

The structure of autonomy, as I defended it in the previous chapter, can be depicted as something like this:



As the diagram above illustrates, to be autonomous is to be powerful and authoritative over one's life. There are certain preconditions of psychological competency - a sense of self, instrumental rationality and so on - required for power and authority, but the status of actually being powerful and authoritative

(in other words, autonomous) is constituted by standing in relations of non-domination whilst possessing these competencies. When we are non-dominated, we are not vulnerable to arbitrary interference in our lives, and this protection against arbitrary interference is based on our being recognised as agents. The autonomous agent, then, satisfies the psychological preconditions and stands in social relations of non-vulnerability and recognition respect.

In this chapter, there are two main goals. The first is to put some flesh on these structural bones by considering types of relations which constitute autonomy and, conversely, relations which are autonomy-hostile no matter how fervently and authentically the agent desires them. I will focus on those relations which are non-dominating (as noted in the diagram above) and show how they are required for autonomy-constitutive relations.

Providing a substantive picture of relational autonomy serves several purposes. As a simple matter of persuasiveness, it clearly serves my case to be able to point to specific relations and explain how they are constitutive of autonomous status (or more-autonomous status, or maximally-autonomous status, etc.), and allows practical problem cases to be directly addressed. From a broader theoretical standpoint, offering these concrete relations will also feed into my eventual argument that no state institution can produce stable relations of the correct sort, because of the dominating nature of such institutions.

The second task for this chapter follows from the picture of relational autonomy so presented, and is similarly relevant to my later claims about authority. I argue that the account of autonomy-constituting relations we end up with suggests that there is no principled distinction between personal (or private) and political (or public) relations: the common view that there are distinct categories of ‘personal’ and ‘political’ autonomy is therefore a misconception.¹ There are certainly kinds of relations which are inappropriate subjects of the

¹ Otsuka (2001), for example, speaks of friendship and romantic relationships as giving rise to “the sorts of injustice which...are rightfully regarded as irrelevant, or only very indirectly relevant, to political philosophy”. Raz, despite rather unhelpfully claiming that “it would be tedious to engage in lengthy argument over what action is political” (1986, p3), clearly thinks that there is some distinction between personal and political - otherwise it would be rather odd to engage in the effort to find a political morality. This thought is supported by his talk of personal and (distinctly) political freedom; see Raz (ibid), particularly pp3-4 and 394-95. Similarly, Tomasi (2004) distinguishes between ethical and political liberalism at least partly on the basis that the latter does not concern itself with “personal moral value”.

kinds of general debate or coercion associated with political decisions, but this is because of the effect that such debate would have on agent autonomy, rather than owing to any necessarily and distinctively ‘apolitical’ feature.² Rejecting the division between personal and political, one of the sacred tenets of liberalism, gets us closer to being able to specify the kinds of conditions that must be met in order for rough personal-political equality to obtain - a necessary precondition for *any* kind of institutional obligation, as I shall argue. We will be able to show that our ‘personal’ autonomy is intimately bound up with our ‘political’ autonomy, and that merely formal equality at the ballot box is not and cannot be enough to legitimate a putatively authoritative institution if material inequalities persist.

The principle that the personal is political (and vice-versa) also allows us to ward off objections that autonomy-based justifications for institutional obligation must be uselessly limited in scope and/or at risk of perpetuating existent social domination; so that, say, the anarchist account of something like participatory democracy fails to consider the effect of informal power hierarchies. If, as I suggest, the only cogent use of ‘personal’ and ‘political’ is as loose descriptors of the extent to which we will sanction public oversight of certain matters, then there is no occasion when we can simply ignore considerations of (for example) inequality or regard them as external to the matter of obligation.

In short, arguing that the personal is still political is to argue that there is no magical protected realm where we may necessarily disregard the political context of our personal lives. Politics and political engagement are, necessarily, woven into every facet of our lives rather than being preserve of some specially-selected elite, and it makes as little sense to speak of personal autonomy without political power and authority as it does to imagine that the brainwashed or entirely subservient agent makes authentic political decisions. As such, a plausible picture of ‘personal’ autonomy is also a plausible picture of ‘political’

² To give a brief example, it is generally accepted that there is some more-or-less stringent limit to which we may interfere with parent-child relations (largely, as I shall argue, because to ignore this limit is to ignore much of why we may value parenting to begin with). Nevertheless, child-rearing clearly *is* a topic of political debate and public input; the question of how I may raise “my” child *is* one that Jane Doe of Anywheresville can legitimately opine on and make normative claims about.

autonomy.

1. Substantiating Substantive Independence: What Kind of Relations Constitute Autonomy?

In the previous chapter, I made the case that only a thorough-going externalist account of autonomy, one which sees the autonomous agent as one who fulfils certain psychological preconditions *and* stands in appropriate relations, can evade the show-stopping objections levelled against alternative accounts. In particular, I claimed that cases like the voluntary slave are archetypal examples of relational autonomy's advantage over its competitors insofar as they show that some kinds of external factors - in particular, some kinds of social relations - are necessarily inconsistent with an agent's being powerful and authoritative over their lives.

The task now, then, is to explain in more detail what relations are constitutive of or hostile to autonomy, and on what grounds.³ After identifying two properties which non-dominating relations must have, and showing how these properties allow non-dominating relations to make us powerful and authoritative, I provide examples of particular autonomy-constitutive and autonomy-hostile relations. The examples should not be taken as exhaustive, nor as a full account of the relations which fall out of a conceptual analysis of relationally-conceived autonomy. Rather, I offer them as paradigm cases, and to suggest the necessary and sufficient conditions for a relation's meeting the requirements given for autonomy. That is to say, the examples are illustrative of what it is for a relation or set of relations to partly constitute our having power and authority over our lives.

It seems appropriate to repeat that the account I offer is similar and heavily indebted to that put forward by Oshana (2006); I rehearse it here both to provide a clear statement of my position and to draw out the distinctions between mine and Oshana's versions of the theory. A point worth re-emphasising

³ I use "autonomy-hostile" rather than the probably more accurate, but rather odd-sounding, "heteronomy-constitutive".

is that the following features are features of the relations themselves; it is not the case that standing in (say) a relation of mutual support *causes* non-domination, but that it *constitutes* it - part of what it *is* to be free of domination, or what it is to be free of some particular form of domination, is to stand in such-and-such a mutually-supportive relation.

After laying out the general idea behind non-domination in more detail, I give two properties that non-dominating relations must have: non-vulnerability and recognition respect. Although the two properties are not simply equivalent to power and authority, I suggest that there is a fairly strong correlation: non-vulnerability is the main player in power over one's life, while recognition respect is key to authority.

1.1. *Non-domination: A First Pass.*

I have borrowed Pettit's term non-domination in order to highlight that our being powerful over our lives does not require us to be "masters of the universe" in order to be autonomous: if it were simply the case that autonomy was a function of being the most powerful and being *regarded* as the most powerful, then feudal lords or monarchs would be the archetypal autonomous agents.⁴

While Pettit's idea of non-domination as a (structural rather than merely coincidental) freedom from *arbitrary* interference with one's life greatly informs the conception I put forward here, there are several relevant differences.⁵ A closely-regulated structure of interference with one's life, according to well-known rules grounded in theoretically-consistent bases, may prove hostile to one's autonomy, but it is clearly not *arbitrary*.⁶

4 The question of whether a feudal lord is more autonomous than, say, a contemporary working or middle-class white Scottish male is I think an open one at this stage: intuitively, the lord has significantly more power and authority in certain realms (as it were), but fails to stand in some significant relations - such as access to reliable healthcare - which more often obtain for the latter. We might thus say that the lord is more autarchic, rather than more autonomous. I return to this discussion in 'Non-domination (2)' later in the chapter.

5 For a fuller picture of his account of non-domination, see Pettit (1996). The difference between structural and coincidental freedom from unjust interference will be discussed further later in this chapter, but Pettit's example of the slave who is a good judge of the slaveholder's whims (and thus avoids some of the slaveholder's worst excesses) is a useful illustration of why structural freedom is required. See pp22-24 of Pettit (ibid).

6 Unless, perhaps, we mean arbitrary in the sense used when describing some action or theory

On my view, an agent is non-dominated if she stands in relations such that she can “determine how she shall live in a context of at least minimal social and psychological security”.⁷ That is, she cannot be constantly looking over her shoulder for death squads or roving bandits (or forced into banditry herself), nor can she be under overwhelming pressure to live in a particular fashion lest she be ostracised, suffer material harms, and so on. The connection between power, authority and non-domination can be characterised at this stage in the following way:

1. *Power*. Without what I term social security (here referring to the existence of accessible systems of universal public welfare, education and so on) my power to achieve many goals is drastically reduced if not simply removed. It will certainly, for example, be considerably more difficult to reliably feed myself in the absence of social co-ordination; and it is hard to see how I could be a teacher, writer, or publican *at all* in a Hobbesian omnipresent state of war.

There are certainly differences between these cases, in the sense that there is a difference between lacking the power to fulfil a role because the role does not exist and lacking the power because you cannot (for whatever reason) access that role; it may also be that some powers are necessarily social in a way that others are not. I probably *could*, assuming that I was a competent survivalist, retain the power to feed myself on a desert island, but it just seems incoherent to say that I could be a teacher in such a scenario - who, after all, would I teach?

It does not seem to me, though, that these differences are relevant to the matter at hand except insofar as we should note the importance of the social to the purportedly personal. It makes no sense to think of autonomy in the outdated terms of self-mastery and non-interference when many of our powers

as making morally arbitrary distinctions, where the arbitrariness is more akin to fixing on morally irrelevant factors than “just plumping” - if non-domination as freedom from arbitrary interference is understood as freedom from *morally* arbitrary interference, there may be less of a division between mine and Pettit’s accounts. To the best of my understanding, however, Pettit holds the view that arbitrariness is something like interference which fails to meet certain democratic criteria; these two claims are not obviously incompatible, but they are obviously not equivalent (on p56 of Pettit (1997), he defends the claim that arbitrariness is something to do with whether or not an act of interference tracks “the welfare and world-view of the public”, but has since sharpened this to include the tracking being done through suitably-formed institutions). See, e.g., Pettit (1997, 1999).

7 Oshana, op. cit., p86.

are *conceptually* dependent upon society, but powers can come in different forms. Sometimes our power to do something is entirely dependent upon there being a social role or function for that thing, as with teaching, and that sometimes there is a socially-independent role which society can help or hinder us in performing, as with feeding ourselves. That there are these different types of power does not, however, make any difference to the claim that being socially secure partly constitutes our having power over our lives.

In terms of psychological security, we can add in the associated worries and stresses of life in a Mad Max-type world. Plausibly, however, the psychological factors are not generally so crucial for power as they are for authority - at least in terms of non-domination. Consider: if for some reason I am psychologically incapable of forward planning or instrumental reasoning, then my capacity for autonomy suffers and it seems odd to speak of dominating relations as undermining my *power* to direct. On the other hand - and to restate a point from the last chapter - we can easily imagine an instrumentally-reasoning, forward-planning agent who nonetheless regards his actions as not following from his self-conception or deeply-held principles. That is, the agent would (may) still have power, but seems to lack authority, under conditions of psychological insecurity.

2. *Authority*. There is probably some sense in which social security *is* required for a person to have authority over her life, as I shall suggest shortly, but we can fairly easily imagine a situation in which some group is blamelessly deprived of basic requirements for welfare (shipwrecks, natural disasters, and so on), and so it would be too strong to suggest that any and all cases where there is a lack of social security entail that agents are not held to be authoritative. Psychological security should be understood as being generally able to plan and carry out courses of action on the basis that they cohere with one's goals, without fearing physical or emotional reprisal.

Note that this does not require that all our plans be successful, nor even lacking in risk; I fairly patently risk reprisal if I plan to assault the next group of strangers to pass by my office, and this does not make me less autonomous. The requirement is instead that we are regarded (and regard ourselves) as

inappropriate subjects of arbitrary coercion - there should be no standing reason to believe that acting on my wishes is likely to result in others harming me.⁸ The reasoning behind this is fairly straightforward: to regard someone as authoritative over their lives is to regard them as an appropriate source of values, and regarding someone as an appropriate source of values is (generally) inconsistent with systematically preventing them from following these values.

The idea behind non-domination as a constituent of autonomy, then, is that the power and authority requirements are met by those relations which obtain between agents/groups of agents and institutions, or agents/groups and other agents/groups (rather than, for example, between the agent's psychological history and reflective capacities).⁹

There are important worries that might well arise at this point regarding the potentially lopsided nature of non-domination. One can, we might think, be non-dominated by living within a suitably utopian, equal society, but surely one can also be secure against domination simply by being the meanest bastard in the playground. Defending an autonomy that lets in 'non-domination' of this kind is inimical to my project. If non-domination only requires that no agent is in a position to interfere with one's life, then whether the interference is arbitrary is of no import - and so it looks as if we're back to talking about non-domination as a feature of autarchy, not autonomy. If, on the other hand, non-domination is concerned with, specifically, *arbitrary* interference, then it will have to be based on something other than brute power. Here's why: the account of autonomy that I argue for includes (as will be shown) something like an internal equality condition. In order for us to stand in relations of maximal autonomy, those relations must guarantee every other relevantly similar agent the same power and authority with regards to own their self-direction. If an agent's non-vulnerability to arbitrary interference is a result of their simply being more

8 And, of course, the opposite belief - that harm will likely result from attempts to self-determine - is partly constitutive of psychological *insecurity* (here meant in my limited sense, rather than the widely-used sense of someone who is 'sensitive' with regards to some trait or other).

9 I refer to "groups of agents" here not in order to begin a discussion of group rights, but to acknowledge that there seem to be times I stand in certain relations only insofar as I am (*de dicto*) a member of some group - considered as the individual James Humphries, for example, I am less likely than many to be assaulted by the police, but as "domestic extremist #32 in George Square" the likelihood is much greater.

powerful than others, then this equality condition is not met (as should be obvious from the fact that such an agent is in the position of being able to arbitrarily interfere with others' lives). I address the concerns about non-domination and inequality in much more detail later, but before doing so it will be worthwhile to expand on some instances of non-dominating relations.

I will therefore now turn to a category of relations which are apt for both domination and non-domination, elaborating on the relevant differences along the way: labour relations.

As our particular instance of labour relations, take the relations which obtain between an agent and their working institution.¹⁰ That is to say, the bundle of relations which obtain between an agent, Amy, her fellow workers, and any other agents who may be in a position to direct or impact upon Amy's activities with regards to her job.

Amy stands in dominating labour relations if, among other things, she is liable to arbitrary or unjust dismissal or punishment; if, for example, she can lose a job as a teacher because the head of school takes a sudden dislike to her.¹¹ Similarly, if she risks dismissal (or punishment, or ostracism, etc.) for refusal to take part in unannounced extra duties, whether or not she does so refuse, she is dominated. The normal way of thinking about having powers to exercise self-direction is that we have the power to x when we can also decide not to x: for example, my being able to throw and catch a ball would not generally be thought of as a normatively important power - or its exercise an important part of my self-direction or conception - if I exercise it only when someone holds a gun to my head and tells me to do so or die.¹² For just the same reasons that a right to work must entail a right not to work, and a right to live correspond with

10 I avoid the putatively more straightforward "agent and their employer" for reasons which will shortly become obvious if they are not already.

11 Bearing in mind my earlier remarks about arbitrariness, it is worth clarifying what I mean by morally arbitrary here. It could be the case that instead of risking losing her job if the head "just happens" to take a dislike to her, she risks losing her job on the basis of some long-held religious bias of the head's - here there seems to be less of the "on a whim" aspect of the first case, but Amy is patently still being dominated.

12 It would perhaps be possible for someone who held a stringently (and highly implausible) Hobbesian view whereby we are free to do anything we're not physically prevented from doing to suggest that my ball-throwing is indeed an exercise in normatively-important powers; this is not the place for that discussion, but suffice is to say that I find the claim unattractive.

a right to die, we must say that an agent's power over their life is marked not just by what they may choose to do, but what they may choose *not* to do. My standing in relations that allow me to teach cannot make me autonomous if those same relations prevent me from doing anything *but* teaching (or unjustly sanction me for not doing so). More broadly, labour relations that coercively tie one to a particular job or life on pain of unacceptable harm or difficulty do not only make one's undertaking of that job heteronomous, but interfere with other powers over one's life that are equally required for maximal autonomy.¹³

With respect to the second requirement for autonomy, Amy cannot be said to be authoritative over her work if her ability to do it - or, rather, the recognition of her ability to do it - is dependent upon irrelevant factors. For her to be authoritative with regards to teaching, it is *not* necessary that "if she says she can do it, she can". We should instead cast the authority condition as being satisfied in the following way. Amy is authoritative over her desire to be a teacher; this authority requires that she be granted the appropriate opportunities and support. If, for whatever reason, she cannot meet the generally-accepted criteria for successful teachership, then refusal to employ her will not amount to refusing to regard her as authoritative over her life. What is required for her to be authoritative is that her desire to be a teacher, *when accompanied by her qualifications to do so*, is an institutionally and inter-subjectively respected and supported reason for her to be a teacher. To summarise the point: if someone can x and wants to x, they are authoritative over their working life insofar as their desire and ability to x is regarded as a good reason for them to x (if they can x but do not wish to, they are authoritative insofar as they need not x).

Dominating labour relations, then, undercut or at the least make unstable an agent's power over her working life, while they fail to acknowledge that agent's

¹³ This is not to say that a dangerous job need make one heteronomous: while it might be the case that economic coercion makes (say) crews of crab boats undertake significant risks of death on pain of poverty, it doesn't seem true that folk will never authentically take dangerous jobs, nor that the labour relations one would then stand in need be dominating or fail to display recognition respect. Rather, the heteronomy consists in standing in relations such that one is coerced into doing a dangerous job *or* that one is coerced into doing an alienating, tedious, etc. (but not dangerous) job, through the threat of harm, poverty or similar. See also Chapter 12 of Cohen (1989).

authority. But what of non-dominating relations?

Amy stands in non-dominating relations if there is nobody with the power to morally arbitrarily or unjustly dismiss her from teaching or sanction her; that is, so long as her power to continue her employment is not dependent upon the whims or brute influence of others. This can be cashed out in ways roughly analogous to the standard negative/positive liberty distinction. Most obviously, Amy is non-dominated when it is structurally difficult for, say, her employer to directly exercise arbitrary power, but she is also non-dominated if there are labour and welfare laws which mean that quitting or being fired from a job need not be a threat to her well-being. To use the usual slogans, Amy is here free *to* challenge or change her employer in addition to being free *from* arbitrary institutional interference; the non-domination consists in her being able to cease a job she dislikes, as well as her being protected against being treated certain ways within that job.

This refers back to my earlier point about coercive labour relations tying one to a job on threat of harm, poverty etc.: even employment which does not directly inflict arbitrary interference can damage our power and authority over our lives if we have to hold that job - however much we despise it - or be unable to pay bills, risk eviction, and so on. An agent cannot be said to be maximally powerful over their life if sheer self-preservation demands engaging in alienating work, and similarly, it is inconsistent with their being authoritative that their only protection against such a situation is brute luck (of having one or more socio-economic privileges or managing to find a job they *do* enjoy, for example) rather than social or institutional support.¹⁴

A couple of fairly minor observations should be made before concluding this section. Although I have generally cast conflicts in terms of employee versus employer, it is entirely possible that one agent (or group) could dominate another within the workplace but without any formal or institutional power to do so. Coercing or manipulating a colleague into doing more than their fair share

¹⁴ This is the point I anticipated several pages ago: social security is required for authority in the sense that an agent can rightly regard themselves as authoritative, but still fail to be authoritative, if their relations are inconsistent with being treated as an independent source of values (i.e. If they are at risk of harm should they give up a disvalued job).

of work, or bullying them (from any number of motivations from personal dislike to oppressive norms) such that they want to quit, are two examples where bosses are at most indirectly involved, but labour domination still occurs. Again, the fundamental problem is the structure of labour relations that allows this to happen. Just as a punitive contract is no less dominating if one happens to have an easygoing immediate superior, the fact that one's colleagues are not reactionary throwbacks does not make being powerless against them if they were any less problematic for autonomy.

In terms of non-dominating relations, meanwhile, I have suggested that trade unions are part of securing non-domination. This should not however be taken as an argument that membership of a (reformist) trade union is enough to attain this security within a hierarchical and capitalist society. Apart from anything else, contemporary reformist unions rely on the means of production being held in private hands in order to maintain their own position; a re-ordering of structures which makes (for example) legal recognition of strikes irrelevant puts thousands of bureaucrats out of a job. There are none-too-subtle implications for the shape of an autonomy-maximising society here, but for the moment I simply make the point that, conceptually hostile to maximal autonomy though reformist unions are, they - in many cases at least - still make their members more powerful over their own lives by acting as a partial counterbalance to the dominating force of employers.

To conclude, then: in this section, I have argued that relations which constitute autonomy must be non-dominating, and given a sketch of what such relations might look like. Non-dominating relations make the agent powerful and authoritative over her life, for example by defending her against the whims of a more economically powerful agent or institution. But unanswered questions remain. I have not provided an account of what *makes* these relations non-dominating (or to put it another way, I haven't picked out the conceptual features of a "defended against economic exploitation" relation which are necessarily shared by other non-dominating relations), nor given any indication as to how we are non-dominated by a secure working environment but not (for example) by being a powerful feudal lord. In the following sections, I explain the

two properties which are individually necessary and jointly sufficient for a relation or set of relations to be non-dominating and thus partly constitutive of personal-political autonomy: non-vulnerability and recognition respect.

1.2. Non-domination (1) - Non-vulnerability.

In the examples above, one noticeable difference between Amy's circumstances under domination and non-domination is the extent to which she is vulnerable to others: whether or not the head of school is in fact fickle or prejudiced - or whether this fickleness or prejudice will as it happens affect Amy - she stands in relations such that she is structurally exposed to the possibility of arbitrary and unjust harm, interference with her self-direction, and so on. Being autonomous cannot be a matter of happenstance or good fortune; the autonomous agent must have a more secure ground for their self-direction than the contingent fact that no-one happens to feel like oppressing them just at the moment.¹⁵

When she is not dominated, however, this vulnerability is not present. There certainly still seems to be the possibility that someone will attempt to harm her or unjustly interfere with her self-direction, but she can strike back, so to speak - she stands in relations such that an unjust attack on her power and authority over herself can be defended against. Take a case where Amy is a trade union member, and will not have to rely solely on her own determination (or wealth) in an employment dispute. There is obviously still the possibility - indeed, the overwhelming probability - that her employer will attempt to coerce her in some or other fashion, but insofar as it is more difficult for them to do so, Amy is less vulnerable, and thus less dominated.¹⁶

Non-vulnerability satisfies the power requirement of autonomy in just the same way that someone would have greater power to build a house if they could do so without fear of attack than if they had to attempt the construction under a

15 As Oshana (2006, pp83-84) says, "we cannot claim that a person is self-governing if her efforts to determine how she will live would have been thwarted had she tried to act differently with respect to activities relevant to the direction of her life".

16 Equally, of course, if it is well-established within Amy's community that exploitative employment practices will not be tolerated, and will be met with general retribution, the effect seems to be the same or greater. This point will assume much greater significance in the final part of the chapter.

constant hail of bottles, or if they had many people to help them (and reduce the risk of, say, their being fatally injured during the process) as opposed to trying to do it on their own. It is, as the fact of the matter has shown, possible for people to pursue their authentically-preferred career under conditions of domination, but exactly insofar as they have to overcome arbitrary and unjust challenges to their power, they are less autonomous.

This leads to an objection that on my view, folk like Malcolm X, Mary Barbour and Frederick Douglass - indeed, all those who stand by their principles in the face of coercion - are less autonomous in virtue of being oppressed.¹⁷ Three non-exclusive responses are possible. We might, first, accept that this seems counter-intuitive but assert that it's nonetheless true: risking death for one's principles is courageous, admirable (at least in this case) and so on, but it is not obvious that this makes one better able to control one's life. Second, we could respond that insofar as they *do* fight off this coercion in order to lead a life more consistent with their self-conception, they are more autonomous (they have lost some powers but gained others which are more vital to them). This is somewhat ad hoc - although not obviously wrong, since the claim isn't that fighting oppression makes you less powerful, but rather that having to do so means that you've started from a less powerful position. Like the first response, though, such a reply might fail to capture the force of the objection, which I take to be that an agent who sticks to their guns *is* in some meaningful sense more in control of their life than the agent in a gilded cage. The final response is to re-emphasise the division between moral and personal-political autonomy described in Chapter 2. The inhabitants of Schier's "Kantian Gulag" might, in fact, be able to stick to all their moral principles, act only on suitable reasons, etc., but does the agent who refuses to nark have more *power* over their lives than the stoolie? Surely not - the stoolie, after all, had just as much metaphysical freedom to assent or dissent, and need not necessarily be irrational or psychologically

17 The objection was first raised to me by Chris Reid during a seminar at the University of Glasgow, and variations of it can also be found in, among others, Meyers (2000) and Sperry (2013), though it is not entirely clear to me whether Meyers is rejecting substantive accounts on this basis, or making the related but distinct claim that agents under oppression are better-placed to assess their own autonomous status than others are to assess it. For further criticisms of (specifically) constitutively relational autonomy on this basis, see Christman (2009), Meyers (2008) and Noggle (2011); for a more detailed response, see Oshana (2014).

fractured.

Unless we are to understand personal-political autonomy as being self-mastery or total integrity - and that, essentially, is what this claim suggests; only the moral hero, spitting defiance at the firing squad, is *truly* autonomous - then sticking to your guns does not constitute being powerful over the direction of your life. While there might well be ways to cast the revolutionary's fight as ultimately leading to or instantiating some kinds of personal-political autonomy, it is misguided to suggest that the revolutionary is more autonomous in virtue of having *less* power over their lives.

In any case, non-vulnerability is important for power not just in preventing oppression or domination by others, but in establishing or maintaining the kind of social structures that allow people to perform fundamentally co-operative tasks, or indeed to co-operate in the first place. Take the Mad Max world mentioned earlier. Whether or not we are unfortunate enough to encounter the Great Humungous (the primary antagonist of the second film), we are in some sense vulnerable just to the vagaries of the world itself - I might be able to teach only when there are no lethal dust storms, or have my power to be a road warrior curtailed by an acute shortage of V8 Interceptors. While there are no agents engaged in oppression, I nevertheless lack powers crucial to my ability to self-direct. Vulnerability can, then, be the result of structural human action, or it might result from disaster conditions. This suggests, intuitively correctly, that an agent is less autonomous (because more vulnerable, and thus less powerful) in times of chronic scarcity.¹⁸

Of course, the possibility of disaster is ever-present - particularly given the omnipresence of dangerous ideologues with nuclear weapons - so it cannot be the case that we are non-vulnerable iff there is a cast-iron guarantee that there will be enough food on the table, warm enough clothes on our back, etc. for the rest of our lives (or that will be no asteroid strikes, nuclear disasters, and so forth). As a fairly rough-and-ready criterion, then, we are non-vulnerable iff we

¹⁸ Of course - as in Mad Max, or with contemporary climate change - these disaster conditions might themselves be human-caused. This probably has a different moral upshot with regards to responsibility and restitution, but from the agent's point of view the effects on autonomy are (likely to be) much the same.

stand in relations that defend us against arbitrary and unjust harm or interference *and* no conditions of disaster/chronic scarcity obtain or are imminently likely to obtain.¹⁹

There is a more pressing objection to field, however. The thought that we might have our self-direction seriously interfered with by the theoretical possibility of an asteroid strike or nanomachine is so far-fetched that it may justifiably be regarded as fundamentally unimportant for practical purposes, but there is one very obvious set of relations in which we seem not just to be vulnerable, but for the vulnerability to be both necessary for the relations and a necessary part of what makes them valuable; namely, intimate personal relationships.²⁰ Since this challenge is one which cuts right to the heart of my overall concern with autonomy and society, it is worth discussing it in some detail.²¹

The objection would take something like the following form: I have claimed that we are vulnerable insofar as our self-direction is prone or susceptible to unjust or arbitrary interference - paradigmatically, if we would (for example) have to give up a valuable activity or suffer serious emotional or physical harm at the hands of another. But this seems to be *precisely* the kind of problem that affects intimately-related agents with divergent or conflicting values; if I am faced with the prospect of passing up valued employment because the alternative is the breakdown of an intimate relationship, then it looks very like I must choose between self-direction and serious emotional harm. As such, my power is compromised; presuming that there is no emotional blackmail going on - that is, I would suffer the harm just as an inevitable result of the relationship ending, rather than because my partner(s) would deliberately inflict suffering on me in retribution - then it doesn't seem correct to say my authority is undermined, but the net result still appears to be that I am less autonomous.

19 As an interesting if bleak aside, this implies that whether one believes humanity as a whole to be vulnerable is largely dependent on how just how close one takes Earth's environmental (and/or capitalism's economic and social) collapse to be.

20 I do not distinguish here between intimate friendship relations, romantic, parental, etc.; there are obviously different types of behaviour that constitute these relations, but the essential vulnerability, what we might unscientifically call the potential for heartbreak or betrayal, is the same across all of them.

21 I focus on the broader question of domination and intimacy in Humphries (forthcoming), where much of this material is developed further.

This presents a direct and troubling challenge to my attempt to provide a thoroughly social account of autonomy: if we are vulnerable, thus less powerful and less autonomous, just in virtue of being in intimate relationships, then other people are a clear threat to autonomy even in run-of-the-mill, non-disaster circumstances, and the kinds of substantive relations that constitute autonomy must turn out to be those which allow no unasked-for limitations of our choice. But this is more or less just autarchy: an individualist, almost Randian and assuredly anti-social view of autonomy as being compromised by the presence of or attachment to others. In other words, we end up right back at the mountain man's door. What is required is a way to show either that vulnerability of this sort is relevantly different to, say, the vulnerability to being bullied at work, or that it is not in fact vulnerability at all.

First, it is important not to confuse our vulnerability to emotional pain with what I have defined as vulnerability. I am, in the ordinary language sense, vulnerable to a swift stab in the eye, but this doesn't mean that I am vulnerable in any autonomy-relevant way. Similarly, we ought not to allow the usual connotations of 'vulnerable' with regards to emotion to colour our view here: anyone in a committed relationship, plausibly, is more vulnerable to heartbreak than anyone who is not, but this should not simply be taken as equivalent to their lacking non-vulnerability in my sense.

We cannot require that deep or committed relationships be ever-lasting; it can't just be the possibility that they end which has the effect (or appears to have the effect) of making us problematically vulnerable. Rather, it seems to be in the yoking of our well-being to that of another. The self-direction threatening vulnerability that I spoke of earlier must, it would appear, lie in the fact we might be at risk of fairly serious psychological damage if we self-direct in certain ways, and - unlike in the case of workplace vulnerability - there is no institution or procedure to watch over such relationships; I cannot launch an emotional tribunal for unfair dismissal.

Here, it will be helpful to consider a point made by Jay Wallace in his paper on obligations of intimate relationships: there are always normative considerations which will result in agents' being "not completely free to chart their course

through life on their own terms” (2012, p189). That is to say, there is invariably the potential for there to be moral or pragmatic reasons not to pursue some course of action we might deeply desire to pursue - but this is not in itself a limitation on our autonomy. My being (morally and perhaps legally) bound not to kick older folk in the kneecaps does not represent an arbitrary or unjust interference with self-direction, so in order for a relationship’s generating obligations upon me to make me vulnerable in my sense, the obligations must be unjust or arbitrary - and with this thought in hand, the anti-Randian path opens up.

Take the plausible thought that there are going to be some obligations which are constitutive of a loving relationship - for example, the obligation to provide support and comfort when the significant other(s) are unhappy.²² If this is correct (and I assume that it is from here on out), and if it is the case that forming loving relationships is at least one type of relation that can be constitutive of autonomy, then we must conclude that these kinds of obligations are, perhaps indirectly, themselves autonomy-constitutive. This surely isn’t a wild claim: I have new obligations as a postgraduate that I didn’t have as an undergraduate, but - at least insofar as these obligations are a required part of “postgraduate-dom” and continued study is part of my valued self-direction - these obligations make up part of my (incompletely) autonomous life. This claim seems fairly innocuous to me.

Note, here, that there’s at least *some* potential for psychological (rather than material) vulnerability; I would, presumably, be deeply unhappy to fail my PhD. But without the possibility of failure, the success would be meaningless, and we can stipulate that this failure is not the result of unjust or arbitrary interference in my life, so it now looks as if the possibility for emotional damage is one which is, at least in my case, a *required part* of the relation. It would be sad, unpleasant, sub-optimal etc. for me to fail, but *not* unjust or arbitrary, so the necessary risk of pain is not vulnerability in the relevant sense. If I have to choose between, say, giving up a valued job offer or giving up a valued

²² I can, I think, remain neutral on whether this obligation is a derivative or, as Wallace would have it, non-derivative and *sui generis* duty of love - whatever its ultimate ground, we can surely agree that being bound by a duty of comfort is just part of what it *is* to love someone.

relationship, either option is going to cause me pain, but then that's surely true by virtue of my valuing them - an exclusive choice between two valued options *must* involve some regret, it seems to me - and it is absurd to think that the autonomous agent is by definition able to obtain everything they value. A painful choice *may* be the result of autonomy-hostile coercion, but it may also be a blameless result of our having relatively complicated and sometimes self-contradictory sets of desires, and in the latter case the appropriate response is somewhere between sympathy and a shrugged "so it goes".²³

The story so far, then, is that neither the generation of obligations by nor the possibility of emotional harm resulting from particular intimate relations will make these relations of vulnerability, at least absent injustice or arbitrariness. Our ruggedly individualist critic is therefore left with just one option to press the claim that committed relationships may render us vulnerable and thus non-autonomous. This is the notion that, in some autonomy-undermining way, we do not freely enter into the varied obligations generated by deep relationships. It is, as the objection might run, as if I *thought* I was becoming a primary school teacher but have instead been manoeuvred into a job teaching at secondary school - or, more to the point, as if I leapt into either job without any idea what obligations might be involved, and am now bound by duties and risks I could not have been expected to foresee.

To draw again on Wallace's paper, the duties of relationships do not simply fall upon us wholesale when we pass some specific checkpoint; "it is partly up to individuals to determine for themselves the exact contours of the obligations they fall under in so far as they participate with each other in relationships of love" (op. cit., p189). If my partner has explained to me that they deem it vital for me to *x*, then either I can reject *x* - with whatever consequences that might entail - or I can accept the obligation to *x*. Now, of course, if the consequences of my rejecting *x* include being, say, savagely beaten or publically humiliated,

²³ The corollary of Whitman's famous "do I contradict myself? Very well then, I contradict myself; I am large, and contain multitudes", after all, is that contradictory multitudes sometimes escalate to conflicting multitudes; and no-one enjoys being caught in the middle of a running street-fight.

then I am vulnerable - but I am vulnerable precisely because of an injustice that should not be present in a loving relationship.

By analogy, if I am told “teaching y is now a requirement of teaching here; do it or teach elsewhere”, it is significantly less plausible to hold that I stand in relations of vulnerability in the workplace than if I am told “teach y or we’ll break your legs”.²⁴ The parallel extends still further: there is no realistic way to guarantee that I won’t come into work and be assaulted by an enraged student/colleague/administrator, but the relations I stand in are (supposed to be) hostile to that sort of thing, so that such an attack would be a deviation from, and guarded against by, the labour relations I stand in. To put it very briefly: I cannot be *sure* that no-one is going to bean me with a brick when I walk into the office, but I am non-vulnerable if I can be sure that such an action would be seen as unacceptable and entirely inconsistent with the employment relations I stand in, and that I will be supported in attempting to fend off and/or seek restitution from my attacker.

We can now shift our focus back to intimate relationships, with the following articulation: standing in special obligations to another does not make us vulnerable, all possibility of emotional pain *and* limitation of self-direction accounted for, unless these obligations have simply been imposed by fiat - and it seems very odd to say that someone could genuinely love (as opposed to be obsessed with, etc.) someone upon whom they shovel arbitrary or unjust obligations. Similarly, it is possible that someone I am in a deep relationship with will act badly: betray my trust, attempt to manipulate or emotionally blackmail me, etc. But, first, this is to act immorally and in breach of “duties of love” in any case, so we are not vulnerable because we’re in a loving relationship, but because we’re in a *dysfunctional* relationship, or just in a relationship which is experiencing more or less temporary malfunction. Second, the other person(s) is/are in no position to force my acquiescence to continued interference of this kind unless I am already vulnerable.

²⁴ Though by no means entirely implausible, depending on the wider context of the first directive - if there are no other jobs and unemployment results in material harm, for example, then there is much less to distinguish between the two commands in terms of domination.

In order to draw out this second point, return to the example of someone who jumps me with a brick in the office. For all that it will be deeply unpleasant, I am not vulnerable in the relevant sense because, as it were, they only get one shot at me; I don't have to plan my work life around evading their continued assault, and there is no institutional reason why anyone else should take over where they left off. I am under no compunction, in other words, to choose between self-direction and welfare. Similarly, the person who acts badly and causes me emotional damage is not in a position to do it again unless I allow them to - and whilst it's obviously possible that I allow them to because of some psychological hold they have over me, this sort of malevolent interference means I'm already non-autonomous or less autonomous - and so the structural, "built-in" possibility of unjust or arbitrary interference does not exist here.²⁵

To recap my argument: the mere fact that relationships may cause us pain does not render us (relevantly) vulnerable, because - given the kinds of creatures that humans are - we are at risk of pain or suffering in indefinitely many activities. It is true that intimate relationships make us particularly easy targets for emotional suffering, but without that risk we would not get the valued relations, just as we cannot (currently, at least) get the valued experience of mountain-climbing without the potential risk of a fatal accident. Further, that we have new and interesting obligations as a result of intimate relationships is not enough for these relationships to be counted as interfering arbitrarily or unjustly with our self-direction, for just the same reasons that *any* kind of obligation-generating relationship need not be interpreted as constitutive of heteronomy - autarchy and autonomy are not equivalent.

Finally, that someone *could* cause us suffering or pain in an arbitrary or unjust way is not the same thing as our being structurally vulnerable to this injustice (unless there are independent relations which, for example, leave us defenceless against abusive partners); apart from the observation that an abusive relationship is of a different kind than a loving relationship, we can add that the

²⁵ The other obvious problem is one where I am forced to stay in a relationship not because of psychological but material coercion - where I am in the situation experienced by abused partners who literally have nowhere to go, or where leaving the relationship would be social suicide, and so on. Here, again, the problem is not confined to the relationship, but involves the other relations I stand in; the fact that the depicted society will fail to defend me, or will enable abuse, is already enough to fail the non-vulnerability requirement.

potential to be hurt by another is not simply equivalent to that other being in a position of domination. To put it in slightly sloganeering terms, it is impossible to be sure that no-one will ever try to hurt me, but possible to be fairly sure that they won't get to try it twice.

So much, then, for the objection that non-vulnerability requires an unpleasantly a- or anti-social account of intimate relationships. We can be vulnerable in, and dominated as a result of, inter-personal relationships, but the requirements of (non-dysfunctional) intimacy are not in themselves hostile to autonomy. What makes us less powerful and authoritative over our lives is not simply that we might experience harm as a result of particular relations, and certainly not that these relations might result in some of our desires being frustrated: it is that the social structures we are embedded in make us systematically vulnerable to arbitrary interference with our lives. Since a non-dysfunctional intimate relationship (like a non-dysfunctional relationship of any sort) ought by definition to involve standing in social relations which do not make us problematically vulnerable, the relational account of autonomy can thus evade the apparent problem in a way which is entirely consistent with what it says about domination in other areas of life. To be non-vulnerable is not to be the mountain man, but to stand within certain social relations, which may very well include as constituents of our autonomy intimate relations and their attendant obligations.

1.3. Non-domination (2) - Recognition respect.

Even if we can cast non-vulnerability in such a way as to prevent intimate relationships compromising autonomy, however, the problem alluded to in Section 1.1 remains - it is all very well to show that one can be autonomous by standing in such-and-such supportive relations, but without showing that the tyrant or feudal lord is not the most autonomous just because the least vulnerable (and thus least dominated or likely to be dominated), my account will end up looking like a desire-satisfaction theory with some additional handwaving, and be vulnerable to much the same problems. In this section, I provide the second condition for non-domination, one that blocks off the problem of might making right whilst also meeting the authority requirement for

standing in autonomy-constitutive relations: recognition respect.

In addition to non-vulnerability, non-domination is partly constituted by what Darwall (2006) calls “recognition respect”: one’s social status and respect being based on the recognition of one’s agency, rather than of some perceived excellence. The latter, Darwall argues (pp122-126), is “appraisal respect”, a form of respect we must earn and which applies to excellence or merit, while the object of the former is dignity or - crucially - authority, and must be presumed or simply granted on the basis of different qualities from appraisal respect.²⁶ That this recognition respect forms the motivational basis for our standing in secure relations of power and authority is a crucial point, and so I now turn to an examination of recognition respect as a property of non-dominating relations.

There are, firstly, no immediately intuitively obvious cases where I can somehow fail at agency in the same way I can fail at some “excellence”. I may, for example, make a hideously ill-advised decision about whether to go to the bar or to go home, but this is patently still an *act* in the relevant sense; I did not, *contra* some of the more implausible interpretations of Kant, somehow cease to be metaphysically free when I settled upon having another round.²⁷ To grant someone recognition respect as an agent is not something we base on an assessment of their ‘success’ at agency. Instead, recognition respect *directly* regulates the authority (and, to some extent, power) we grant to others; when we offer someone recognition respect, we acknowledge their authority. For an agent to be authoritative over their lives in the sense required for them to be autonomous on my account, they must therefore have recognition respect: if we do not recognise someone as an agent, and hence do not recognise the *prima facie* authority this grants them to determine the course of their lives, then they do not meet the Strong Authority Condition referred to in the previous chapter.

It might be objected here that there is clearly still an evaluative aspect to recognition respect: unless we are to adopt some intuitionist approach whereby I

²⁶ I use Darwall’s terms for convenience’s sake.

²⁷ Or, to put it more modestly and forestall determinist objections, I did not cease to be compelled to act *as if* I were metaphysically free.

“just know” to offer this respect to the appropriate candidates, there must be some feature or cluster of features which is motivating the decision to regard Jones the Neighbour as an authority, but not Jonas the Neighbour’s Dog.^{28 29}

It is true that recognising someone as an agent requires some amount of evaluation - plausibly, *any* kind of recognition requires evaluation of the object of recognition - but this is not the same as that recognition respect being, itself, an evaluative trait. That is to say, the kind of evaluation we undertake to determine that someone *has* a trait that demands recognition respect is not simply equivalent to judging that the trait is present to such-and-such a degree, whereas the objects of appraisal respect can be evaluated both in terms of their ‘mere’ presence or absence and in terms of the extent of that presence.

To give an example, say that you grant me appraisal respect as a philosopher to the extent that I meet particular conditions. If I meet 30% of these conditions, you might grant me 30% respect (to use a weirdly mathematical abstraction). Perhaps, for example, I don’t *always* affirm the consequent, but often do so; this lessens my appraisal respect just insofar as I do it. But I cannot only sometimes be an agent, or only sometimes have basic human needs. If you make the evaluation that I am an agent, then the respect demanded admits of no apportioning according to “how much” of an agent I am.

To pursue the point further, consider the difference between agents and what we might call mere actors; beings such as songbirds, mice, tuna, and so on. There is clearly some level of consciousness or awareness here. Such beings respond to stimuli and, as best as we can tell, have experiences, and there is obviously something driving the mouse to prefer peanut butter over cheese. But the mouse does not suffer when there is only cheese, and has no awareness that some malign force has been replacing all the peanut butter with cheese in an attempt to make a philosophical point. We could plausibly say that it might be wiser, or in the actor’s best interests, for us not to interfere with their goings-on, but it would seem strange to cast this as our interfering with their self-

28 In fact, we plausibly *should* be motivated to regard Jonas the Neighbour’s Dog as some kind of authority, but the knotty question of non-human and non-sapient animal agency is one I must simply leave aside.

29 Carter discusses this problem of evaluation in his paper on equality and respect - see Carter (2011), pp540-543.

determination - to slightly alter my claim above, how can there be self-determination without a self?

An agent, on the other hand, must have some idea of themselves as (conceptually) distinct from and able to influence the world around them. Unlike actors, agents can make choices rather than simply respond to stimuli; whether or not these choices are authentic, well-thought-out or whatever else, they are still the result of a decision-making process (that many or most non-human animals do not seem to possess). It might, thus, be more helpful to reconsider the notions of pseudo- or proto-agents sometimes used to characterise human children's capacity for choice, and instead think of agency as something that can be applied across a greater or lesser (more or less complex, more or less considered, etc.) range rather than something that comes in greater or lesser degrees. Rather than there being many kinds of agency - primate agency, child human agency, adult human agency - determined by conceptual sophistication, I suggest agency as a determinable with many possible determinates; one can either be an agent or not, but what one is an agent *about* may vary wildly.³⁰

There will undoubtedly be hard cases here with regards to, for example, persons with particularly severe kinds of brain damage, but it should be remembered that there are already preconditions for autonomy (say, instrumental rationality) which they may not meet. Whether or not someone possesses the various autonomy competencies is not simply equivalent to whether or not they are an *agent*, and recognition respect simply requires the latter; we cannot be autonomous without recognition respect, but we do not demand recognition respect *on the basis of* our autonomy.³¹ To put it another way: to be heteronomous because of the relations one stands in is a different thing to being heteronomous because of some psychological state or incapability, but the former agent is not thus more morally considerable nor more deserving of recognition respect than the latter.

30 I owe the idea of casting agency as a determinable with many determinates, or multi-tracking disposition, to a discussion with Umut Baysan and John Donaldson at the University of Glasgow.

31 In fact, it seems to me that we cannot demand recognition respect on this basis - how could we be autonomous if autonomy required recognition, but we were owed such respect in virtue of our autonomy?

To summarise: there are fairly broad (and, as I have given them, very sketchy) evaluative properties that we might use to determine agency. Once this determination is made, however, the form of respect demanded - recognition respect - gives prescriptions on how we may act and what sort of reasons we must consider, rather than how we should evaluate the property.

Recognition respect is thus required for non-dominating relations because a non-dominating relation must hold the relevant agent to be authoritative over themselves and recognition respect forms part of agency's normative demands; as such, it allows the relation to satisfy the authority condition of autonomy.³² In practical terms, recognition respect probably makes demands of the power conditions as well, but whether or not this is so might well be dependent on who or what it is that purports to be doing the recognising. It seems odd to say that I grant recognition respect to someone whose attempts to self-direct I consistently arbitrarily or unjustly frustrate, but it also seems that I could give an agent recognition respect while being unable to affect their power over their lives in any way.³³

The idea of recognition respect as playing a normative or prescriptive role leads to the second important point on the matter. Whereas appraising someone as an excellent crab diver (that is, granting them appraisal respect as a crab diver) gives no guidance on how to do anything *except* appraise them as an excellent crab diver, recognising someone as an agent is intimately bound up with treating them a certain way. Namely, recognising someone as an agent requires that we take their desiring something to be a *prima facie* authoritative reason for them to pursue that thing; that is to say, while I can consistently give someone

32 Or perhaps satisfies the relational component of the authority condition.

33 This suggests a further, more complex problem (posed by Oshana): what is the minimum power threshold required for recognition respect, and why isn't meeting this threshold enough for non-domination? I think the most I can do here is to suggest that the 'minimal social security' I referred to on p126 gives a rough guideline for the powers we *must* have in order for it to be plausible that we are recognised: if I am denied food, shelter, healthcare, or am at constant risk of violence, then I am not being recognised as an agent. Similarly, if my wanting to do something is seen as giving *no* good reasons for me to do it, it seems odd to say that my agency is recognised and respected. This is a necessity, not sufficiency condition (as is suggested by the discussion elsewhere in this chapter); I can have enough food and not be looking over my shoulder for bandits and still not be fully or properly recognised. Perhaps, for example, someone agrees that my wanting to be a teacher is a good reason for *me* to do it, but denies that this gives anyone else any reason not to interfere - and so it is possible to meet this minimum threshold without thereby being non-dominated.

appraisal respect as a philosopher while also denying that they have any principled rights to self-determination, I cannot grant them recognition respect and deny them these rights.

This may seem in tension with what I have just said about the possibility of hard cases - surely, the objection comes, we *can* respect an agent but also think that, on the basis of some descriptive fact (or, more cynically, some 'descriptive' 'fact'), they should not self-direct, or not self-direct in such-and-such an arena of action?

Here again it is useful to consider the agency-autonomy relation. One cannot be autonomous without being an agent; it is nonsensical to speak of someone determining their own way of living if there is no "their" in the first place. But one *can* be an agent without autonomy - apart from children or many non-human animals, there are clearly many autonomy-competent agents who are nonetheless not autonomous. To recognise someone as an agent is to grant them authority and presume their rights to autonomy given certain psychological requirements; to be autonomous is to be a competent agent who stands in the appropriate relations. Thus, denying that an agent *in principle* has rights of self-determination is simply incoherent.

Still, the critic has a move to make here. There does not as yet seem to be anything preventing the would-be dominator from performing the argumentative equivalent of the regretful shrug and claiming something like "yes, *in principle* the agent should self-direct, but what about this most disadvantageous feature which *in practice* prevents them from doing so, or justifies interference?". Here, the dominator's claim would be that the agent's *prima facie* authority is recognised but over-ridden by other factors: thanks to poverty, or gender, or ethnicity, or sexuality, or whichever characteristic is seized upon, the agent is ultimately not to be regarded as authoritative over the direction of their lives. There are innumerable historical examples of just such a tactic being deployed to deny autonomy - indeed, basic welfare - to groups so characterised. Without successfully countering it, my theory is irrelevant or worse, providing support for oppression.

One pragmatic response can be made here: as a point of fact, these kind of moves are probably if not overwhelmingly going to be based on palpably false premises or fallacious or bad-faith reasoning. It was manifestly *not* the case, for example, that the stolen generations of Australian Aboriginal children were taken from their parents because of some good-faith, well-reasoned but mistaken assessment about those parents' competence to self-direct, or the children's competence regarding their own wishes. It was done, and knowably done, to destroy community identity and to further secure the dominance of white Australians. My theory, then, does not have to make any special moves to defend against problems of this sort precisely because the problems arise *after* the point of departure from my theory. One pretty obviously cannot interfere with self-direction for the purpose of buttressing racist and colonialist power structures without thereby failing to grant the agent recognition respect, since apart from anything else displaying racism towards someone is clearly not to accord them recognition respect.³⁴

Still and all, some theoretical response is required - we might not be able to make any account rogue-proof, but we are certainly obliged to make it rogue-unfriendly. Evaluations about an agent's autonomy competency may legitimately affect *how* we respect their agency, but they may not legitimise our acting in a way inconsistent with their actual competence. How this actual competence should be determined is a question that would require a separate paper - or many papers - to answer fully, but some fairly general thoughts can be put forward, all of which suggest that requiring recognition respect as part of an autonomous life does not imply an ableist or overly strenuous account of autonomy.

First, and uncontroversially, we ought to err on the side of caution: someone's making a decision that seems odd or even 'crazy' cannot be grounds to think they are incapable of self-direction, although repeated out-of-character or

³⁴ Again, it is worth noting: individuals can display racism (e.g. by using ethnic slurs), but structures can do so also - by, for example, being insensitive to (or contributing to) social conditions that result in some group or groups being disadvantaged in some particular way. The idea of structural racism is of course an old one; I bring it up to emphasise the point that whether the individual, group or structure is thought of as acting primarily in the private or the public arena, the failures of recognition respect occur in precisely the same ways.

wildly inconsistent behaviour of some kinds *might* be a sign that something is psychologically (or, given the plausible thought that many kinds of putative mental illnesses are the predictable results of or coping mechanisms for structural/social conditions, structurally) amiss. To give a blunt example: if I shit on the floor of my house for erotic pleasure then you have no grounds to interfere, no matter how odd it seems to you, but if I shit on the floor to ward off the demons I think live under the carpet, there are probably grounds for investigation - although even then it is not obvious that I am incapable of self-direction. As with many aspects of self-direction, there are likely to be times when it seems absurd to us that any autonomous agent would want to x or y and we are nevertheless bound to let them x or y; the possibility of a non-autonomous act being treated as autonomous is in some circumstances far less worrisome than the reverse.³⁵

If someone is, for whatever reason, not capable of choosing between three options, then it is not overbearing for an appropriate stand-in to choose for them. Indeed, it is plausibly part of a non-dominating society that I may have friends or others who can be trusted to act on my behalf in the event of incapacity. If I am recognised as an agent, I have standing rights for my interests to be considered in a way that clearly isn't true of a rock or a car; that I might temporarily or permanently be incapable of furthering or protecting these interests myself in no way changes this. The evident possibility of hard cases, and of bad faith assessments, cannot be taken as evidence against this claim; people will deny the obtaining of all kinds of relevant factors (or claim the relevance of irrelevant factors) if it suits them, and it does not seem unduly pessimistic to say that no quantity of paper sacrificed to demonstrating their lack of philosophical justification will prevent this from happening.

Secondly, we should consider that *individual* judgements of incompetence are only going to become damaging when we are already vulnerable - I have no reason to care that Colonel Blimp regards my kink as a sign of madness unless

³⁵ The issue of adaptive preferences is particularly pressing in political philosophy generally and leftist politics specifically because of the false consciousness problem - "if you want to x, then that's proof enough that your reflective capacities have been malignly influenced by so-and-so!". I addressed some of these issues in Chapters 2 and 3.

Colonel Blimp has the power to non-trivially interfere with my self-direction. A society or set of institutions organised so that its members stand in autonomy-constitutive relations - that is, non-vulnerable and instantiating recognition respect - must be a society that is not prone to casually interfering with self-direction.

It should be clear now that while there is a psychological aspect to recognition respect (at least in terms of individuals), it is still a fundamentally relational matter. If, to give a historical example, there is putative institutional/legal respect for agents - as with the British government's attitudes to the "Windrush generation" of Caribbean migrants to the UK - but the same agents face violence or hostility if they attempt to claim the rights associated with this respect, then it is simply *not* the case that they are being granted recognition respect.³⁶ To use a familiar concept: purely formal recognition respect is no respect at all. However, a situation whereby one has all the psychological indicators of recognition respect - the community doesn't hold racist, sexist, homophobic, etc. attitudes - will probably be a situation where formal or institutional rights are upheld also. Again, this is primarily a sociological or psychological point in the sense that particular attitudes will *probably* produce particular social structures, but it does seem incoherent that a group of people could both hold attitudes of genuine respect towards me as an agent and also believe that I should be denied associated rights.

To recap this section's key argument, then: recognition respect, as opposed to appraisal respect, is owed to an agent just in virtue of their being an agent, and requires that certain rights (such as self-determination) be presumed. Although this respect does not act as an all-things-considered trump card on all occasions

³⁶ To give a brief illustration: assuming (very unsafely) that having a good chance of putting a representative in Parliament, or becoming one, makes one more politically autonomous, the structural and institutional racism of the 1950s UK militated against African and Caribbean Britons being elected, plausibly making these agents less autonomous. There was no formal bar against persons of colour being elected, but the kinds of attitude encapsulated by the 1964 Tory election campaign for Smethwick, which ran with the infamous strapline "if you want a nigger for a neighbour, vote Labour", were endemic. Whether or not one is legally barred from running for parliament, the public propagation (and *at best* institutional endorsement by toleration) of racist hate speech by a powerful faction within that parliament renders the prospect of winning a slim one, and the prospect of running a potentially life-threatening one. To be in a position whereby attempting to access rights of respect risks predictable, because structurally endemic, violence is clearly not to be respected.

- you do not deny me recognition respect if you ask me to take my briefcase off your foot - it *does* provide *prima facie* authority to the decisions of any agent, an authority that requires good reasons before it can be over-ridden.

There are both psychological and formal or structural requirements for such respect to be instantiated, but it is difficult to cleanly separate these requirements from each other. This, I would argue, is a feature rather than a bug. It would be odd if not incoherent to imagine a world where one is granted psychological but not formal respect, and history shows that formal respect alone cannot make up relations that instantiate recognition respect.³⁷

We can now connect recognition respect to the power and authority conditions in the following way. The relations of recognition respect, requiring as they do that I do not face undue difficulty in self-directing, make me more powerful over my life. If the neighbours do not hold the kind of attitudes that will cause them to throw rocks at me (or approve of rocks being thrown at me) while I garden, and the state or governing institution presumes and protects my right to garden just because, as an agent, I ought to be able to garden I want to, then I am more powerful than in the reverse situation.

The way in which recognition respect fulfils the authority condition should hopefully be fairly clear, since it is more or less built in. If we are regarded as *prima facie* authoritative because of our agency, as recognition respect requires, then we must be regarded as (at least *prima facie*) authoritative over our lives. What is probably the more interesting feature of recognition respect in terms of authority is that this respect - and as such, our authoritative status - is itself instantiated relationally and in large part non-psychologically. I am authoritative not just insofar as my neighbour thinks I'm an appropriate source of values, but also insofar as any changes of heart they might have do not change my authoritative standing in the wider community. If, to emphasise the point, I can be regarded as an inappropriate source of values on the basis of one person's say-so, then it is surely odd to assert that I was truly authoritative to begin with

³⁷ As it happens, the only circumstance I can think of where "psychological but no institutional" respect *isn't* incoherent is one where social attitudes have outpaced legal change; even there the fact that the legal system can lag behind in such a fashion plausibly points to some shortfall in psychological respect, but it at least seems within the bounds of possibility.

- rather, it seems that I was coat-tailing on another's authority, or that I had false or pseudo-authority.³⁸

Taken together, the non-vulnerability and recognition respect conditions of non-dominating relations ensure that we are powerful and authoritative over our lives in the way required for autonomy: we are structurally supported and protected against arbitrary or unjust interference, and supported not for any unstable or inappropriate reasons like privilege or influence, but simply in virtue of our being agents. Note that the relations of non-domination should be understood as constituting autonomy insofar as they constitute primarily *programmatic* autonomy. To dust off the lumberjack example used in the last chapter, non-domination does not consist in our being able to chop down any tree we feel like, but in being able to be a lumberjack at all. Social relations which constitute our being powerful and authoritative over our lives need not (indeed, surely could not) allow us to do anything we want at any given time, but they must allow us to pursue our values and life plans insofar as those values are consistent with maximal equal autonomy.³⁹

Notably, the recognition respect requirement goes some of the way towards blocking the objection that non-domination could be attained just by being the dominator. Even if such an agent is not currently vulnerable - perhaps, for example, because of their political or economic position - this non-vulnerability is based on morally irrelevant features rather than recognition respect, and so they ultimately fail the non-domination condition and are less than maximally autonomous. The recognition respect requirement thus militates towards what I shall call maximal equal autonomy-constituting relations: even if we meet the power conditions for non-domination, if we meet these conditions because we stand in relations of privilege then ultimately we fail the authority condition. We are not respected in virtue of being recognised as agents, but excused from

38 This thought, that what appears to be recognition respect may in fact not be depending upon what its real basis turns out to be, will be returned to in the final chapter.

39 This, again, shouldn't imply that only relations which allow us to pursue suitably 'weighty' or 'serious' life plans constitute autonomy. Although it seems relatively unlikely that anyone would have as their central life goal the collecting of (for example) biscuit tins, if they did, then their programmatic autonomy would be constituted by their standing in relations which allow them - effectively, not just formally - to collect biscuit tins. It would not, though, be harmed by inability to obtain a particular biscuit tin, in the same way that the lumberjack's autonomy is not limited by being unable to cut down a particular tree.

challenge in virtue of being privileged, and this just puts us in the position of prison trustees or favoured servants - enjoying the ride all the way up until the point where the screws are investigated or the boss goes bankrupt and we are thrown under the bus.⁴⁰

This in turn suggests that theories of institutional legitimacy which seek justification in autonomy need not concern themselves with being “value-neutral” - being autonomous is *not* just a matter of deciding on things in a particular way and being able to carry those plans out, and so autonomy-minded justifications are not forced to accept that, for example, exploitative economic practices could be legitimised on the basis that a majority of people authentically voted for them (or, conversely, that a losing minority would have its autonomy reduced by being prevented from exploiting other agents).

This has obvious import for any account of political obligation or authority we might try to develop; before moving into that part of my thesis, however, it is worth drawing out another feature of my theory which has similarly weighty implications.

2. Structures, Relations and Why the Personal is Still Political

There is a final implication to be drawn from the account of autonomy I have given in this and the previous chapter, and one that should serve to link the account with my broader aim of identifying the requirements which institutions must meet in order to be both authoritative and autonomy-respecting or constitutive. The implication is that the relational account of autonomy has no room for a sharp conceptual split between ‘personal’ and ‘political’ autonomy.

40 We can conceive of this as their still standing in dominating relations - as it happens they are (at the moment) the beneficiaries of such relations, but because their position cannot be based on consistent moral respect, they are essentially only autonomous only insofar as they meet some morally arbitrary set of criteria. But, of course, this is the position that the slave in Pettit’s example occupies, and we do not want to say that they are autonomous. Indeed, pushing it up to the top of the hierarchy, we can see that the politician, aristocrat or mob boss whose privilege is maintained by entrenched injustice *still* stand in such relations, for their self-direction is protected via violent exercise of power, not mutual recognition respect; the difference is a matter of degree, albeit great degree, not of kind. Our sympathy is not required, and this is again not to say that they are worse-off than folk who are both vulnerable and not granted recognition respect, but the point that non-vulnerability must itself not be based on arbitrary features, and must reflexively include an equality condition, is worth emphasising.

So far, the requirement that autonomy-constitutive relations include recognition respect and non-vulnerability has included no specification of how these features have come about - as alluded to in the discussion of labour relations, it seems that I am protected against economic domination if I have a powerful trade union *or* if I stand in community relations such that friends and acquaintances will defend me against attempts at exploitation. But - traditionally, at least - trades union activity is in the political or public sphere, while having strong relationships with friends is counted as a matter of personal or private preference.⁴¹ This distinction cannot solely be a procedural matter, either. I am perhaps more able to rely on an institution specifically designed to protect me against exploitation than on friends who may or may not be efficacious, but that itself is unsatisfactory as an explanation of a distinction between personal and political autonomy given that my trade union may be less effective than a suitably protective and motivated group of friends.

Similarly, the way we interact with our friends and relatives is often thought of as being not a matter for political philosophy (or a matter only to the extent that we point at a vague area and say “not our problem!”). This, the thought seems to go, is the moral philosopher’s realm and we ought not to get our political mud all over the floor if at all possible. But treating, say, a trans agent hurtfully on the basis of their identity is not merely immoral, it is an action that (tries) to define the status and the conditions of people of a certain identity within society. That is, surely, an archetypally political matter.

I argue that the personal-political split is a confused one which ought not to be replicated in our conception of autonomy: someone can satisfy what we might think of as formal conditions for political autonomy (no legislative bar to engagement/voting/etc.) and yet still be politically heteronomous. As I shall suggest, the best way to explain this apparent inconsistency is by pointing to exactly those kinds of relations which are taken to be important to - or, on my view, constitutive of - ‘personal’ rather than ‘political’ autonomy.

41 Again, Otsuka (2001) defends the personal/political split; see also Raz (op. cit.), and Oshana (op. cit., p102) for the claim that an agent can be politically autonomous without being personally autonomous (and vice-versa).

There are, in essence, two interlocking claims which make up my argument against the personal-political division.

First, if we understand political autonomy as being limited to interactions with or powers against particular formal or legal institutions, we will be unable to satisfactorily explain why agents like the angel in the house are not able to effectively engage with political institutions. We will also be unable to explain to explain how agents may have extensive powers to effect political change despite being formally disempowered or disenfranchised. In other words, the account of political autonomy as holding formal rights or being able to access certain decision-making procedures casts the net both too widely *and* too narrowly in terms of capturing the character of ‘political’ autonomy.

The second claim is that the power and authority over our lives required to be personally autonomous must include the same kinds of power and authority (or power and authority over the same kinds of areas) that are taken to be crucial to political autonomy: the personal and the political are thus different emphases on the same concept, not two different concepts with a shared foundation.

In this section, the task is therefore to give plausible accounts of the kinds of relation that are picked out by “personal” and “political”; identify the features that are supposedly doing the individuating; and argue that in fact we can find putatively distinctive political features of personal relations (and vice versa) to such an extent that the terms are really only useful to indicate very general areas of public discourse. On the picture which emerges my personal life is personal insofar as it is on the whole none of your business, but the reasons that it is (on the whole) none of your business are overtly politicised, and insofar as my personal life intersects with many others, it is, at least potentially, very much your business.⁴² Political autonomy thus entails personal autonomy, and vice versa. This is a conclusion which is important in providing both a plausible account of autonomy and plausible autonomy justifications for political authority.

⁴² As Williams notes, “To say that “the personal is political” is to say only that private life is implicated in networks of power, leaving open the question of under what circumstances it is appropriate for government to intervene” (2001, p94). I would perhaps use “community” rather than “government”, but the point stands.

2.1 *The Personal and the Political*

The distinction between the personal and the political in terms of autonomy is usually characterised in something like the following way: personal autonomy is a matter of doing what one wants (within certain limits), of being free and able to settle on the direction of one's life and work towards the values and goals that one has. Political autonomy, on the other hand, is seen to be the holding of certain rights against one's government and/or fellow citizens, specifically with regards to involvement in (or being involved by) coercive institutions.⁴³ We can be politically autonomous without being personally autonomous - so the agent who is not self-directed, or whose desires are inauthentic (etc.) can still be politically autonomous insofar as, say, they have certain legal rights or protections with regards to voting, taxation, and so on - and by the same token we may be personally autonomous without being politically autonomous: a standard example would be the disenfranchised agent who happens to be so far off the grid that they are entirely free to pursue their self-conception.⁴⁴

Both forms or realms of autonomy are taken to have at least some shared foundations - in both cases, we hold *prima facie* rights against interference/of assistance just by virtue of being agents - but the one is not usually explicable by reference to the other. Although we might claim that at least some measure of political autonomy is necessary for personal autonomy, the reverse claim is far less common. The existence of that hoary chestnut, the benevolent dictator, is a marker of the belief that personal autonomy is (at least potentially) achievable under conditions of fairly extensive political repression.⁴⁵

The divide is not merely one of emphasis, either - at least, in the standard liberal narrative - for as Lee says, "The idea of a limited government is

43 Wolff (1998) defines political autonomy using this criterion.

44 Note that in the latter case, they're only personally autonomous if autonomy doesn't require things like access to public healthcare, education, or indeed the ballot box, which should provide an early indication that the distinction is built on sand.

45 Of course, there are some cases where, even on the standard view, political heteronomy is inconsistent with personal autonomy: if part of my self-conception is "non-trivially engaged with the institutions that affect much of my life", then being prevented from so engaging will be counted as making me less personally autonomous. But this is a contingent, not conceptual, link from the personal to the political - I'm only less personally autonomous because, as it happens, my personal autonomy is partly reliant on my political status, and (at least on non-relational models) if I didn't care about political status I would be autonomous without any change in that status.

predicated on the assumption that it is both possible and desirable to keep politics out of the private sphere”; if there is no sharp conceptual divide between personal and political then it would be misguided to suggest that politics had no business in the home (or that familial relations were an inappropriate subject of social or institutional oversight).⁴⁶

Perhaps because of its status as a presupposition of many theories, it is difficult to find a recent account of the personal-political division that gives necessary or sufficient conditions for something’s being personal or political. Instead of a straight counterexample to such putative conditions, I will instead consider a case illustrating what I take to be the core notions of the personal or private sphere.

Imagine an agent deciding where to go for dinner tonight. Here, they are (personally) autonomous insofar as they stand in relations such that they can go to a range of friends’ houses or restaurants, they are competent instrumental reasoners, will not be bullied into picking one kind of food over another, etc. They are, albeit in a fairly minimal way, powerful and authoritative with regards to their self-direction. What, though, are the properties that make this a specifically personal autonomy? Three possible contenders present themselves: intention or goal, decision-making process, and scope.

Taking the intention approach will give us something like the following line of argument. Political action (and therefore political autonomy) is concerned with a particular set of intentions or aims, specifically those to do with the broader society one lives in. While this encompasses our desires to some extent - we have desires exactly about the nature of our society - the relation is only one-way; that I have *some* plan does not make my carrying that plan out a political action unless the plan is aimed at special kind of goal or good. Here, picking a place to eat tonight is a personal action because it is silent about social goods or institutional organisation; one cannot read off a political statement from a claim about what kind of food I’d like for dinner.⁴⁷

46 Lee (2007, p164).

47 The obvious objection here, of course, is that deciding what to eat - insofar as it may be a decision about the value of non-human life - is a political decision even on the standard model. I agree entirely, but let us pretend for the sake of argument that this angle is

Putting the emphasis on decision-making, meanwhile, provides a different account. Here, the thought is that realms of action might be divided up (at least partly) by what kind of reasoning or evaluative procedure we use to settle on a particular course, and - more pertinently - what kind of procedure we *ought* to use. Thus, while I can answer the question “where should I eat” with reference only to my own tastes and preferences, questions like “what sort of businesses ought to be allowed/supported/banned/etc.?” must include input from others, and input of a specific (if broad) type - that is, input through something like a democratic ballot or consultation.

Lastly, making the argument that personal and political action are individuated through the scope of the action results in a story whereby matters of personal autonomy are limited in scope to oneself and one’s friends/family, while political activity ranges, potentially, across national and global institutions. There is some overlap with the intention account here, but they are not equivalent. For example, perhaps I want sole use of some patch of nearby land, without a care in the world for any social or institutional implications. While I do not *aim* at any obviously political course of action here, the scope of such an outcome (affecting as it does any members of the community who might have claims on that land) move it out of the personal, and into the political domain.⁴⁸ When I muse on where to eat, though, there is a very narrow realm of action to choose within; perhaps one person or group has to sit at a different table or go to a different restaurant, but this is not the same thing as changing the ownership of means of production (or recreation) as in the field example.

2.2 *The Personal is (Still) Political*

We thus have three ways of explaining the intuition that decisions about dinner are personal rather than political. However, I now argue that none of these ways give us any distinctive criteria for what is political about (for example) a decision on whether or not to take strike action. That is: there are no sharp conceptual differences in intention, decision-making processes or scope between

irrelevant (or imagine that I am deciding between solely vegetarian or carnivorous eateries in this scenario).

⁴⁸ Alford and Hibbing, to give an example, seem to have the scope distinction in mind when they claim that “political temperament [as opposed to “interpersonal” temperament] deals with the structure and organisation of large-scale social life” (2007, p197).

the two cases, and we should conclude that ‘personal’ and ‘political’ at most mark out areas where certain different types of interference (or coercion) may or may not be permissible.

Take the aforementioned example of deciding whether to strike. What conceptual distinctions can we draw between this and the restaurant decision in terms of intention or goal? Obviously the immediate intention is different (in the way that the immediate intentions of all distinguishable actions are different), but this doesn’t do any useful work: the interesting question is whether the important features of the intentions are different - whether replacing “where should I go for dinner?” with “should I cross that picket line?” requires radically changing the kinds of reasons which bear on the answer - and it seems to me that they aren’t. When I decide where I want to go for dinner, I am making a claim about how I want my life to go (albeit briefly), about what is important to me and what kinds of actions best fit with my ideals or principles.

We can say *exactly* the same thing about striking: I want my life to go better, so I strike for better pay (working conditions, etc.); I claim that labour power is important to me, or that sufficient time with my family is important to me, so I strike; I have ideals of fair working conditions, or collective action, or economic justice, which are inconsistent with crossing a picket line, so I strike. Here, someone could object that this is inconsistent with what I entertained earlier about the idea that deciding where to go for dinner is a personal rather than political decision. The mere fact that both dinner and striking involve plans, reflect my goals, etc., cannot be enough to put them both in the same category. This is true, but it misses important facts about the extent to which both plans presume and require certain background conditions. When I decide that I want to go to place A rather than place B for dinner, I am making a decision which is affected by (and in turn affects) the social relations I find myself in. If, for example, A is more expensive than B, but I prefer it because it is known to pay its staff better, then I am making a decision about self-interest versus social values - just as I may personally lose out as a result of strike action and still decide to strike. *Contra* the earlier claim, we *can* read off political ideals from dinner decisions - it is just a bit more involved than reading off such ideals from decisions about whether or not to go on strike.

This won't, however, put the objection to rest. The interlocutor could still assert that, absent the kind of self-conscious moralising which I seem to be engaged in, the dinner decision does not reflect political views - I just want to have something tasty to eat, and damn all other considerations.

My response is simple: I just deny that such a thought would be apolitical. Take the strike contrast case again, but change the thought process so that I *am* solely concerned with my own immediate benefit, and stipulate that in this instance I will lose out by going on strike (whether through lost pay, being marked down as a troublemaker by management, etc.). Pretty obviously, I will not strike - I just want to earn a wage, and damn all other considerations. There are many ways we might describe this agent: selfish, short-sighted, and so on. But would we say that they have acted apolitically? Plausibly, no, because the presuppositions of their decision are manifestly concerned with questions of solidarity, the role of self-interest in economic relations, and so on. We cannot simply divorce "what matters to me" from "what matters to others", nor from how it affects others: denying that the decision to scab is a political one *already presumes* a sharp division between the personal and the political, and presumes that something must involve a vote, or a parliamentary body, or what have you, in order to be political.

To put it another way, this objection presumes that there are such things as purely self-regarding actions, and this simply isn't so for any action taken in a social context.⁴⁹ If we follow it to its logical conclusion, the criticism also seems to imply that actions and decisions which involve others - such as "where should my family go for dinner?" - would be political *just* in virtue of not being self-regarding, which I take it that nobody wants to endorse regardless of their views on the personal/political distinction (or the possibility of solely self-regarding actions).

⁴⁹ Two qualifications should be made here. First, as I have implied, there *is* conceptual space for self-regarding actions - if I am the last agent on Earth, or trapped on a desert island with no hope of rescue, then plausibly questions like "should I sleep under this or that tree?" really *are* solely self-regarding. Most of us, however, are not in such a position. Second, what we might think of as mental actions (including affective responses as well as propositional attitudes, etc.) might well be self-regarding; but this is orthogonal to the problem and hence a question I need not be concerned with.

To the second possibility, then. The claim here is that dinner decisions are personal, and strike decisions political, because of the kinds of procedure we undertake to make them or because of the kinds of reasons that are inputs to our decision-making process. Thus dinner decisions are personal because the correct method of determining where I want to eat is simply to ask myself where I want to eat. Again, though, this line of criticism rests on dubious presuppositions. If we imagine that the question of strike action is to be put to a collective, then there obviously *are* procedural differences - but if we put the question of dinner to a collective, we (ought to) follow a different procedure than if the question was solely answerable by us. If, on the other hand, the question remains “ought I to go on strike?”, then we cannot claim that the eventual decision will or will not be political depending on whether I canvassed other people’s opinions.

One factor which might cloud our reasoning here is that oftentimes, the procedural requirements for making some political decision *will* include (say) voting or some other method of collective input, and this makes it appear as if anything which doesn’t have such requirements must therefore be apolitical. We should be careful not to put the cart before the horse here: procedural requirements are set by the nature of the decisions being made, rather than such requirements being definitional of what kind of decision it is. When I consider whether I ought to go on strike, then I am making a one-person political decision. When a collective considers whether it ought to require its members to go on strike, then it is making a collective political decision, and must use procedures appropriate to collective decision-making. But the reasons which motivate this (for example, considerations of equal input) remain the same across both cases; all that changes is the particular procedure those reasons motivate us to adopt.

To sum up my response to the claim that personal and political are individuated by the kinds of decision-making procedures we adopt: there are probably reasons to think that individual and collective actions (or decisions) should have different procedural requirements for legitimacy, but believing that this split maps onto personal and political only follows from an implausible presumption that an individual’s decision is personal and a collective decision is political. We

ought to take a vote on where we collectively go for dinner for the same kinds of reasons that we ought to take a vote on whether we collectively strike, namely that our autonomy is plausibly compromised by having no opportunity to affect and determine collective decisions. The relational structures are parallel, although not identical. By the same token, while my individual decision on some matter - whether it be dinner or striking - ought not to be a matter for general debate because deciding what *I* think is a key precondition of being autonomous, this motivation gives no distinction between 'personal' and 'political' decisions.

Finally, there is the third possibility: the personal and the political can be distinguished by the scope of the action or intended outcome. This criterion would account for why procedural distinctions arise *after* the definition rather than forming it: actions with broader scopes require a broader spectrum of inputs. So when I decide on my dinner, there is no need for a vote because the scope of the decision is very limited (its effects range only across myself and whatever cafes or restaurants I am considering), while when we deliberate on whether to strike we do need a vote because the scope of the action encompasses many other individuals, groups, and institutions. This *does* suggest a rule of thumb for distinguishing between those decisions which are primarily individual and those which are primarily collective. When I decide where to eat, I make very limited relational changes if I make any at all, but strike action might (at least in theory) change a much wider array of social and economic relations. But note that such a rule makes no distinctions of *kind*, only of degree; the reason to elicit a wider range of input for 'political' decisions is because the scope ranges more widely, not because political decisions fall within a unique category which invariably demands collective decisions.

I have argued here that conceiving of autonomy relationally suggests that there is no principled distinction between the personal and the political: my being able to choose between two restaurants is not just indicative of but also dependent on a whole host of factors which are typically characterised as 'political', such as the economic relations we stand in and our practical ability to self-direct according to our own principles. Similarly, deciding whether or not to take part in strike action involves or may involve considering the effect that it will have on my family, or on my relationships with friends - the political

decision is also a personal one both because of the factors relevant to the decision and because the decision may alter the personal relationships that we have. Neither intention nor procedural requirements can separate out the personal from the political, and different scopes of decision or action do not require a difference in kind between the two putative types of autonomy.

I do not claim that *only* the relational account implies the conclusion that the personal is political (such a claim would plainly be ignorant of history, apart from anything else), but I do claim that conceiving of autonomy as a matter primarily of our relational standing entails that any alterations to that standing must be of the same broad category, the category of “decisions or actions which alter the relations we stand in”. We cannot be personally autonomous and politically heteronomous, nor vice-versa; only personally-politically autonomous to a greater or lesser extent and perhaps unevenly across different areas of our lives.

2.3 An Objection to Collapsing the Personal/Political Distinction

Before concluding the chapter, there is one further objection to address. Rather than take issue with any of my critiques of putative distinctions, this objection claims that collapsing the distinction opens up our lives to unpalatable levels of oversight and interference from others. The problem can be summarised like this: if everything I do is political, then why is it not a matter for public discussion who I sleep with, or make friends with, or have nothing to do with?

A response to this objection is already discernible within what I have written so far, but it would be helpful to make it explicit. There is no *a priori* reason why matters like intimate relationships should not be the focus of public discussion, but there *are* strong autonomy-minded reasons to think that this kind of social setup would be unfavourable. Remember that we are concerned with autonomy as a matter of being embedded in social relations which are consistent with maximal equal autonomy. When we consider actions which would change those relations for a large number of people, we should (at least) get the input of those people, since we are non-trivially affecting their self-direction - though we may still think, for reasons given in Chapter 3 and earlier in this chapter, that their self-direction can be over-ridden. They get oversight and input, in short,

because to do otherwise would be to deprive them of powers which are required for maximal equal autonomy. But if we give a public assembly powers to interfere with my intimate or romantic life, what have they gained - how do these powers allow them to more effectively self-direct in a fashion consistent with others having maximal equal autonomy? They have gained nothing, and I - by dint of being unable to significantly determine how my life goes in this respect - have lost considerably. This does *not* mean that there can be no general norms or social rules governing relationships in the abstract: the impermissibility of sexual violence does not prevent me from self-directing in a way consistent with others' maximal equal autonomy, but a general norm that nobody could go out for a drink with anybody else without a quorum *would* prevent me from directing in such a fashion.

Another way of looking at it can be illuminated with this analogy: there are good autonomy-minded reasons for a community not to permit the trading of certain products - let us say teabags which are produced by slave labour - because their exchange would result in the endorsement and, if they're sold, financial reward of practices incompatible with maximal equal autonomy. But there are, *ceteris paribus*, no good reasons for my tea preferences to undergo the same level of scrutiny because we would get nothing and lose much; practices of letting people choose their own teabags from within a morally acceptable range are, surely, more conducive to maximal equal autonomy than everybody's having to have their breakfast drink okayed by a committee.

We have, then, a claim that the maximisation of (programmatic) autonomy provides its own set of rules with regards to others' being able to scrutinise and interfere with my life. If the scope of some decision or action ranges widely, it is very likely that autonomy gives us a reason to take the decision a certain way and on the basis of certain inputs - for example, by taking a vote on the action from any relevant agents. If the scope is very limited, we have a reason to take the decision a different way and on the basis of different inputs - so that my decisions over whether to eat *here* or *there*, which in total form a not-insignificant part of being able to decide how my life goes, ought to be taken with reference to where I want to eat and very little else. But this doesn't make the former decision political and the latter personal; it means just that one is

apt for collective oversight (in the way *usually associated* with the ‘political’) and the other is not. When I decide to prefer a Celtic pub over a Rangers pub, that is a political decision; it is also a personal decision about where I want to spend the next couple of hours of free time.⁵⁰ The personal really is (still) political, and the political is personal; for this reason, I will refer hereafter to personal-political autonomy rather than distinguishing between two concepts.

Conclusions

I have argued in this chapter that autonomy-constitutive relations, in order to make us powerful and authoritative over our lives, must make us non-vulnerable to arbitrary interference with those lives, and must be based on recognition respect. That is, we must have a *prima facie* right to do what we want with our lives (ultimately conditional on our doing so being compatible with maximal equal autonomy), and this right must be based on the acknowledgement of our status as an agent deserving of recognition respect.

If we are constantly vulnerable to interference with our lives, then we are not powerful over them, for being powerful over one’s life must mean something more than its direction not *currently* being interfered with (this also means that effective rather than merely formal power is required for autonomy). Similarly, being able to self-direct because it pleases somebody else to allow us, or happens to align with their own goals, is not to be authoritative over one’s life. If my being a teacher is solely permitted (or encouraged) because it furthers somebody else’s aims, rather than because my status as an agent gives me the right to pursue being a teacher, then the direction of my life is not explicable by reference to respect for me as an agent and hence I am not authoritative over how my life goes.

⁵⁰ As in the example of tea produced by slave labour given previously, if it was to transpire that the reason I preferred some Celtic pub was because it was a racist establishment and only served whites then we would have autonomy-based reasons not to let the pub exist in the first place (or to address the relations which allowed such a pub to exist and survive, and so on); notice that this isn’t the same thing as saying that my decision to go there or not ought in principle to be overseen by some committee of public morality.

Finally, I have argued that there is no meaningful normative distinction to be found between those relations which make us ‘personally’ autonomous and those which make us ‘politically’ autonomous. They both, ultimately, concern our ability to direct our lives and change our material conditions in accordance with our values and self-conceptions, and the only plausible difference is one of scope (and thus one of degree rather than kind). This does not, I’ve claimed, make us vulnerable to oppressive levels of interference in our lives because the same considerations of autonomy which collapse the personal/political distinction mean that some regions of programmatic autonomy - such as those concerning our intimate lives and our everyday preferences - are best served by individual rather than collective decision-making.

In the next chapter, I return at last to the problem which opened this thesis: reconciling an account of autonomy which (though social) values self-direction, with an account of authority which must include the ability to over-ride at least some self-direction and, on my view, will sometimes do so *on the basis of autonomy’s value*. With the relational conception of autonomy now in hand, I argue that a notion of authority as functioning to promote and defend maximal equal autonomy can solve the problem of political obligation.

5.

Authority (I)

Introduction

After concluding the first half of this thesis, we have an account of autonomy which starts from the claim that autonomy is a matter of being powerful and authoritative over our lives. This status of power and authority, I have claimed, is largely constituted by standing in particular relations, and has conceptual limits set by the nature of our demands for recognition. By this I mean that valuing autonomy, rather than just valuing the satisfaction of our desires, requires that we value autonomy for others just as much as for ourselves. Even if I authentically desire a life where I go about setting fire to people, for example, I cannot justify it on the basis of autonomy because such an attempt would be internally inconsistent with a commitment to autonomy: valuing relations which make agents powerful and authoritative over their lives is not compatible with valuing relations which allow folk to be set alight at another's whims.

The aim of this chapter is twofold: to “turn the ship” from defending an account of autonomy towards constructing an anarchist account of political authority, and to indicate as clearly as possible the problems of voluntariness and universality this course change will encounter.¹ Having explained the problems here, the next chapter will (attempt to!) solve them.

First, I will recap the theory of autonomy defended in the first half of this thesis, rehearsing a couple of paradigm cases and emphasising the external requirements for autonomy, and sketch the problems they seem to pose to political authority. Then, I lay out desiderata for our theory of authority; I suggest three conditions for this autonomy-motivated concept of authority, and go into more detail about what the conditions require and why they are required. Finally, I consider two conceptions of political authority, the Razian

¹ Here, and everywhere else in this chapter, I use “political” in the thin sense of being appropriate for collective decision-making or to do with collective co-ordination, rather than to suggest something conceptually divorced from the personal or the private.

“service conception” and Klosko’s authority based on obligations of fairness; I claim that although neither conception will do the job required by our concept of political authority, there are useful features in both and the reasons for rejecting each conception provide further signposts for what our eventual account of political authority should look like.

One note before continuing: there are two separate concepts which are picked out by “authority”, and it is important not to confuse them. On the one hand, there is the authority over oneself which is a foundational component of autonomy: the normative claim that I am best placed to decide for myself what I want to do, and that my wanting to do something is a *prima facie* reason for doing it. This is the authority that I have been concerned with in my chapters on autonomy.

On the other hand, there is authority as a characteristic of power structures; something which plays a role in the interaction of autonomous agents within social structures. It is something like the right to provide and have heeded some reason for an authoritative (in the autonomous sense) agent to do something which they have decided for themselves not to do, or vice-versa. The two kinds of authority are thus in principle compatible at least if we understand them correctly: something’s being authoritative over me in the second sense need not make me non-authoritative in the first sense. The problem is not that the existence of external authorities makes me non-authoritative (over myself) by definition, in other words, but how to structure and justify these authorities in a way that recognises and is consistent with our caring about agents’ authority over themselves in the first place.

1. Autonomy Recapped, and the Problem of Authority

The autonomous agent, I have argued, is authoritative and powerful over her own life: able, albeit within certain limits, to decide the course of her life and have this decision respected and supported in virtue of institutional respect for her agency (i.e. recognition respect).

These power and authority requirements are satisfied in two distinct ways: psychologically or internally as preconditions *for* autonomy, and relationally or externally as constituents *of* autonomy.² I may, for example, stand in social relations that would enable or facilitate my power over my life, but be incapable of practical reasoning. Conversely, it is also easy to imagine an autonomy-competent agent who stands in relations that deny their power and authority. It will probably be helpful to provide some cases of these contrasting failures.

The first type of autonomy failure can be illustrated with the following example. Take some ‘average’ adult of a rigorous democracy, who stands in relations such that they would be supported in their plan of (e.g.) becoming a teacher, and who would be supported in virtue of it being their goal. That is, they are acknowledged as an appropriate source of values with regards to what they want to do with their life, and empowered to pursue those values through access and opportunity to take up training courses, material support during periods of unpaid training, and so on. *Ceteris paribus*, they are authoritative and fairly powerful over their lives.

But now let us add that all else is not equal: the agent also suffers from mental ill-health which has the effect (among others) of making them psychologically incapable or barely capable of actually following through on their plan. We thus have a situation where (by hypothesis) they value being a teacher; where this valuing is reason enough for institutions to support their pursuit of becoming a teacher; but they cannot exercise the powers to engage in that pursuit.³ They

-
- 2 Friedman describes these preconditions as competencies, and draws the distinction “between, on the one hand, conditions that constitute the autonomy of choices and actions, and, on the other hand, conditions that are causally necessary for the realisation of autonomous choices and actions” (2003, pp11-13). Friedman holds something like a procedurally-independent desire-satisfaction account with relational elements, hence the talk of *choices* constituting autonomy and external factors as merely effective on the realisation of that autonomy: the parallel in my account would be the preconditions for autonomy and the constitution of autonomy, but they occupy roughly the same conceptual space. The distinction can also be thought of with this analogy. In order for me to be able to play football, I have to meet certain requirements: some capacity for forward planning, minimal bodily co-ordination and so on. If these requirements are met, I stand ready to play football - but it is the running about, kicking a football, sliding through opposing players etc. that actually constitute the playing of football. As with football, so with autonomy.
- 3 I have discussed the possibility of some psychological problems directly undermining one’s *authority* in Chapter 4, and will not return to it here; I stipulate just that in this case, the

are thus less autonomous not by dint of external oppression or interference, but through failing the preconditions for autonomy: the race is not rigged, but they have not made it to the starting line.

The second kind of autonomy failure, and the kind my thesis so far has been most concerned with, is illustrated by simply reversing some of the previous example's claims. As before, the agent values becoming a teacher. Now, however, they are in a psychological state that presents no barriers to their pursuing this value, but they stand in relations that make them powerless to pursue it, or they would be supported in the pursuit only for reasons other than their valuing it. Examples of this kind of heteronomy would include folk who are (whether formally or substantively) prevented from training to be a teacher, or who are supported during their training only because and insofar as they are shaping the educational narrative in accordance with the wishes of the powerful.

Here, we can imagine that the agent is at the starting line, eyes locked on the finish - and as soon as the gun goes off they are tackled to the ground. Although there are, as discussed, plenty of cases where external conditions can *lead* to an agent's eventually failing autonomy preconditions, it can never be the case that an agent stands in autonomy-hostile relations whilst also being autonomous, whereas an agent may stand in autonomy-constitutive relations while failing the precondition requirements.

To re-emphasise the point: I have argued that the relational components of autonomy can both affect the meeting of preconditions and (if they instead constitute heteronomy) make those preconditions indeterminate of whether an agent actually is autonomous; the relational constituents, rather than the preconditions, should therefore be regarded as the primary concern for autonomy-minded political philosophers.

As I have argued in the preceding chapters, this account captures many of our intuitive and theoretical platitudes about autonomy. That there are some

problems are primarily action- rather than assessment-undermining - I know what I want to do, I just cannot motivate myself to do it (or to figure out how to go about doing it, etc.). Severe depression seems to me to be an example, although obviously not an "all and only" type example, of how mental ill-health may undermine the prerequisites for autonomy.

occasions when we can be told what to do (or what not to do) without this constituting heteronomy is thus unsurprising in at least one way. To recall the example that opened this chapter, the account will claim that I cannot kick my heels and bawl “repression!” if I’m hauled off a passer-by and the lighter forcibly removed from me as I attempt to pursue my deeply-held goal of setting fire to unsuspecting pedestrians. This claim gels neatly with our thoughts about the role of autonomy within moral as well as political philosophy: we cannot, on the basis of commitment to autonomy, demand exceptions for ourselves.⁴

The flat-footed example above is not, however, the only instance of what we might think of as blameless denial of self-direction. Indeed, it is plausibly the least problematic instance. According to the schema I have laid out over the preceding chapters, the only reason we might think otherwise is through a misconception of autonomy or a failure to fully work through the implications of some act or other. If, in other words, I am prevented from carrying out an autonomy-hostile act, or my attempt to justify some act by reference to autonomy is unsuccessful, then the mere restriction of action is no threat to autonomy under the relational account I have put forward. Autonomy properly construed is not simply autarchy, and to reduce options or restrict actions is thus not necessarily autonomy-hostile. This much, at least, we can defend against the mountain man’s demands for autonomy-as-autarchy.

The more serious problem arises when we attempt to reconcile this picture of autonomy with another of this thesis’ goals. I aim to provide an account of political authority that takes due heed of autonomy without collapsing into conceptually sleek but practically useless total voluntarism; a voluntarism whereby I can be obliged - and perhaps coerced - to do (or refrain from) only those things that I have explicitly contracted or agreed to do (or refrain from). Under this picture, there can be no institutional obligation to give aid to the

4 As I have argued in Chapters 3 and 4, commitment to autonomy “but just for me” is ultimately self-defeating. Not (or not just) in a Kantian way, but in a material consequences sense: if I want folks’ autonomy-constituting relations to be vulnerable to the whims of a would-be oppressor, then I either already stand in, or am committed to standing in, relations that make my autonomy a matter of luck. That is, if I am (more) autonomous because I occupy a position of privilege, then my autonomy is dependent on this arbitrary privilege rather than on my being respected as an agent - consequently, I am still vulnerable and thus not maximally autonomous.

person who collapses outside my front door, nor - unless I have agreed to be part of it - any obligation to contribute to a public health service.

And this, of course, is exactly what the mountain man asserts: only through his explicit consent can he be placed under (or, perhaps, can he place himself under) political obligation. To borrow from Simmons' (1979) summary of the problem, my account of autonomy gets us the voluntariness desired for political obligation, but looks as if it will have great trouble with universality. If everyone has to agree, then there will always be some folk who we have no claim over; but if we're willing to coerce non-volunteers, then why profess to value consent as a marker of self-direction?

Now, if I were putting forward some free market libertarian political theory, this would be a selling-point (as it were). Only freely-accepted contracts or promises allow external coercion; so much the better, says Nozick (1974) and, although he seems understandably reluctant to acknowledge it, Wolff (1998) probably ought to agree. However, political anarchism has a historical - and in my view correct - hostility towards hyper-individualised theories of obligation that sits ill with accepting that *any* authoritative command is a limitation on autonomy or freedom. I explore this problem in more detail in the next chapter, but in short, the worry is that "against all authority" is ultimately somewhat unilluminating for anyone who cares seriously about how an anarchist society might look.

In addition to the problem of reconciling voluntarism with the ambition of universal political obligation, there is the question of what any putative authority might be entitled to enforce. Often, we take the view that person A may legitimately demand that B stop doing something, but may not intervene to *force* the cessation. For example, it seems reasonable to ask the person behind me on the bus to turn their music down, and if we not implausibly assume that blasting unsociably loud music on public transport is a (very minor) wrong then I am in some sense making an authoritative demand; nevertheless, forcibly turning off or confiscating their phone would not be permissible.⁵ Either there

5 An interesting example of precisely this situation can be seen in *Star Trek IV: The Voyage Home*, when Spock uses the Vulcan nerve pinch to render a stereo-toting punk unconscious - he is then applauded by a busful of passengers who are apparently fine with what is

can be the possibility of ‘empty’ authority - “you ought not do that, but I ought not stop you” - or, more plausibly, different kinds or extents of authority justify different kinds or extent of coercive acts.

This is a somewhat separate issue, insofar as “what justifies authority?” and “what does authority (or being authoritative) justify?” clearly admit of different answers, and arguably different *types* of answers, but it seems very likely that our answer to the question of how we should think of authority will shape what, if any, coercion that authority may legitimately employ. To return to the bus example, my authority as a citizen or fellow traveller may not extend to confiscating possessions - but perhaps my authority as a busconductor (or whatever) *does*. Thus, one way of sharpening what we might mean by political authority is by examining what kinds of action are consistent with the conceptual shape and basis of that authority.

2. Desiderata for an Account of Authority

So much for autonomy. If, as I have argued, the correct conception of the autonomous agent is of an agent who is powerful and authoritative over her life (where these features are partly internal but primarily externally constituted by social relations), then we have half the puzzle of political institutions solved - we know at least one value that these institutions should promote (evince, respect, etc.), and we know roughly the kinds of relations that would tend to be constitutive of or hostile to this value.

This still, however, leaves a significant challenge to tackle. In order for any institution or group of institutions to actually go about promoting this value, then we must know what powers they may use to this end. Otherwise, we have a target but no rules or guidelines on how to attain it. By analogy, if we were to tell someone that football exists to promote the scoring of goals, and tell them to play football, and nothing else, we could not be surprised if they picked up the ball, hacked down opponents and so forth. Understanding *why* we value

effectively assault, so long as the victim is kind of obnoxious.

autonomy (football, etc.) gives us some hints as to what may or may not be carried out in its name, but a political philosophy with pretensions to practical usage must do more than gesture at the likely shape of the institutions it is concerned with. There are thus two questions to answer in order to lay out some rules for authoritative institutions: in what realms do we want them to be authoritative, and for what reason?

I take these questions in reverse order. The desiderata, or what we want authorities to be authoritative over, must surely follow from the reasons we need a concept of authority in the first place.

2.1 Implications of the Functional Analysis

Recall that in Chapter 2, I argued for a functional analysis of authority: the content of the concept is given by what purpose we want that concept to play (rather than being a pre-existing concept with clear necessity and sufficiency conditions which we must fit into our theory). There, I suggested that the functional analysis allowed us to explain *why* it was required for an authority to have the right to command (and be obeyed), and what would justify such a right in the context of political institutions, more easily than the conceptual analysis.

So much for why we should perform a functional analysis of authority. The question remaining is what we would lack in the political state of nature, to borrow Lane's phrase again - that is, what we would lack if we had no concept of political authority.⁶ It certainly seems that communication, co-operation and the like would be possible, so the task now is to give an autonomy-minded explanation of why we would need to create the concept.

To put the issue in the starkest terms possible: what reason do we have for surrendering or exchanging our autarchy to an institution which may make and enforce decisions we disagree with? This is not the worry that anything save unrestricted choice is heteronomy - my argument doesn't imply the conclusion that having obligations of care to a loved one makes us less autonomous, as I have shown - but rather that we seem to be rendering ourselves vulnerable to

⁶ See Lane (1999).

commands that have a justificatory shape alarmingly similar to “because I said so”. Arbitrary interference with one’s life is heteronomy-constitutive, after all: being under the thumb of somebody (or some body) such that the course of your life is determined at their say-so seems a paradigm case of heteronomy.⁷ So what makes institutional obligation *non*-arbitrary? What function does authority play here which justifies authoritative institutions having coercive power?

The Hobbesian answer to this is straightforward: the obligation is arbitrary in the sense that there’s no particular reason for person or body X rather than Y to issue commands, but we have good reasons to obey it anyway because the point, the function, of an authority is to resolve co-ordination problems.⁸ Since, on the Hobbesian view, making co-ordination possible is both the necessity and sufficiency test for authority, anything which *does* do that - in Hobbes’ case and given his view of human psychology, this ability is a direct function of bigness of stick - is thus authoritative and can command us more or less as it pleases.

Perhaps because of the notoriously unappealing methods which Hobbes proposed as tools of proper authority, it is easy to miss the relatively modest and beguiling view of its function that he offers; it is not difficult to spin a much more benign interpretation which leaves us with exactly the same problem.

We might, for example, think that humans generally tend towards the Kropotkinian rather than the Hobbesian caricature, and still think that informal bonds between agents and communities will (just as a result of complexity, human frailty, etc. etc.) be insufficient for the just and progressive resolution of resource disputes.⁹ In other words, it seems entirely coherent to believe that, a) Humans are ineliminably social rather than atomised beings, b) We usually tend towards co-operation rather than competition, and c) The function of government/authoritative institutions is simply to facilitate co-operation.

Despite starting from an almost diametrically opposite position from that of Hobbes, by c) we have ended up giving exactly the same purpose - and thus

7 Here I once more draw on Pettit’s (1997) useful account of domination as vulnerability to arbitrary interference.

8 See Hobbes (1651), particularly Chapter xvii.

9 See, in particular, Kropotkin (1972).

sufficiency criterion - for authority. Perhaps something about a) or b) might motivate us to introduce limitations on how authorities may 'facilitate' co-operation, but it is certainly worth noting that Rousseau's legislating general will has at its disposal very similar powers to Hobbes' sovereign.

It does seem plausible that part of the purpose of political institutions is to resolve otherwise intractable co-ordination problems. If we are deadlocked over the choice of "x or y?" and there are no moral considerations in play, it seems obvious that a body with the ability to decree "x!" will allow greater progress just by virtue of being *something* to move matters along. A community of individuals genuinely committed to maximal equal autonomy, but irreconcilably split about how best to achieve it - even in micro-level cases such as the use of some patch of land - is presumably not going to make much actual headway on creating autonomy-constitutive relations, after all. Hanrahan & Antony (2005, p63) make the point that:

Sometimes decisions must be made in the absence of a rationally sufficient reason to do one thing rather than another. And sometimes, even when an optimal solution exists, the cost of finding it can be prohibitive. In such circumstances, the *best* decision is the decision that *gets made*. Consider, for example, the hell of visiting BlockBuster to choose a movie with your partner. After a while, almost any film seems better than spending another minute in the video store. (This is the only reason that *Gigli* has ever been rented.)

In other words, sometimes authorities must go for A rather than B for no particularly good reason (although I presume that there are good reasons to do A or B in the first place). This being so, we want the decision about A to be binding - it is, as previously mentioned, unsatisfactory for our purposes if the institution selects A, but we shrug our shoulders and carry on B-ing. Notice that an account of moral authority cannot do the job here: if it is equally morally permissible to A or B, then we can do either. Crucially, someone trying to make us A rather than B would be in the wrong - when our actions are morally underdetermined, any action not already ruled out is permissible according to purely moral considerations. So political authorities must be able to do something more than just assert and enforce (even correct) moral claims.

There is a further qualification to be made with regards to seeing political authority solely or primarily as a method of resolving co-ordination problems: there are obvious cases where expediency triumphs even when, intuitively, it ought not to. If the *only* role of political authority is to move resources about and resolve disputes in the most efficient manner possible, then it doesn't matter much why we think that's its only role. Whether King Solomon believed that both plaintiffs deserved to suffer the loss of a child, or that neither would possibly countenance the death of an infant for their own desires, he still threatened to bisect a child. If that kind of swift resolution is the sole purpose of a political authority, it's difficult to dispute the wisdom and, more importantly, authority of Solomon. Clearly, though, this is not what we want of authority: I have already claimed that the task is to find an autonomy-minded justification for authority, or a conception of authority that is compatible with my account of autonomy, and the picture of authority we get from Solomon fails to perform that function.

For the reasons given in the preceding discussion, we should conclude that although resolving co-ordination and resource distribution problems is part of what we desire an authority to do - because without co-ordinating and co-operating it will be nigh impossible to create the kind of social relations that are constituents of autonomy - this cannot be all that we want it to do.

So what, if not (solely) for dealing with co-ordination issues, does authority do - why should we bother with it?

Primarily, I argue, authorities act as markers and promoters of values; we want authoritative institutions to issue rules or decisions that serve some political end. So an egalitarian will identify some institution as authoritative because it issues directives that instantiate and/or promote the value(s) of equality; a market-libertarian authority enacts policies consistent with the ideals of minarchism and is thus authoritative; and so on. Note that this implies not just different justificatory strategies, but different realms of authority, depending on the sought-after values: the liberal-democratic separation of public and private, for example, fits much more neatly with market-libertarian than egalitarian roles for authority. This is hardly a ground-breaking point, but it is nevertheless

worth emphasising the extent to which substantive judgements suffuse allegedly uncontroversial or conventional political presuppositions.

In the case I present, because we are (by hypothesis) committed to relations that constitute maximal equal autonomy for agents, the purpose of authority is to defend or promote these kind of relations. Since we have established that autonomy is not best conceived of as something most easily attained in isolation, then it follows that this authority must extend further than the individual in order to properly perform its task.

Rather, the authority or authoritative command should stand as something like a marker of decisions, institutional policies etc. that will promote or defend autonomy-constituting relations. Since we are, by argument (see earlier chapters) committed to maximal equal autonomy if we care about autonomy at all, then we want some action-guiding principles; thus, the political authority must be able to say “do *this*, because it is most autonomy-promoting policy to follow”.

So a political/institutional authority, whatever it may in fact look like, must provide us with good reasons to think that its policies are consistent with maximal equal autonomy on the basis of recognition respect.

2.2 Desiderata for Political Authority

The previous passage suggests that - in my case, at least - the purpose of authority is to give us autonomy-minded reasons to act, reasons that are consistent with our being coerced into obedience. Here I am mostly concerned with why or how being authoritative in such-and-such a regard is helpful for being able to promote or instantiate autonomy-constitutive relations.

In this section, then, I suggest three desiderata for a theory of institutional authority, and explore what makes them desirable. These desiderata are, in order:

1. The putative authority must be able to generate (morally) binding obligations.

2. These obligations don't stand independently of the institution (for example, "don't murder" isn't a politically authoritative claim, it's a morally authoritative claim that often gets made by political institutions but would still stand regardless of whether any institution asserts it).
3. The justification of and process of enforcing obligations must be consistent with the theory of autonomy I have suggested so far.

1. *Binding Obligations*

The first requirement is reasonably straightforward and, I think, intuitive: political authority must be able to issue binding obligations. That is to say, when some institution tells me to *x*, it should not be a suggestion, or merely an interesting report on their policy, but an imperative to which disobedience is (at least *prima facie*) morally wrong. There are, I think, two major motivations underlying this desideratum.

The first motivation is conceptual. Without a moral upshot to its claims, any institution seems more like a descriptive 'advisory board', and as such inapt for the task of administering or enforcing rules. Consider some equivalent to the General Medical Council, which tells us that eating such-and-such a food, or drinking so much per day, is bad for us. This surely falls victim to a version of the famous Open Question argument, in that it does not seem particularly odd or inappropriate for someone to say "Fine, *x* is bad for me, but why ought I not to *x*?" (perhaps we should say "why, all things considered, ought I not to *x*?").¹⁰ Similarly, I have no reason to disbelieve that Pluto does not meet the criteria for planetary status met by Earth, Mars etc.; nor that paper burns at 451 degrees Fahrenheit, but there is equally no convincing reason for me not to go around saying "Pluto is a planet", "Paper burns at 440 degrees", and so on.

Without falling back to the Weberian idea (2004) that authority is a matter of being the only one acknowledged as a proper wielder of the state's big stick, we surely cannot say that a putatively authoritative body with no moral claim on obedience (rather than, at this stage, the means to enforce it) is doing the work

¹⁰ See Moore (1993) for the original form of this argument.

we require of a political institution. To emphasise the point, political authority ought to give us an imperative that is in some sense conclusive - they say “jump”, and we do wrong if the question is “why?” rather than “how high?”. A political authority cannot generate solely prudential reasons, not least because there are always non-absurd cases where prudential reasons fly in the face of the imperative.¹¹

The second motivation takes the conceptual thought in a slightly more practical direction. For any institution to promote or defend relations that are constitutive of autonomy (or indeed heteronomy) it must have some power. Quite what form this power takes we can leave open for now, but it seems correct to say that while an institutional authority may prove ineffectual in enforcing moral obligations, it must at the very least be able to *try*, if it is to be even in theory able to create or protect autonomy-constituting relations.

A bit more needs to be said about how institutions may or may not generate obligations in this way. It will not be enough just for some institution to assert itself as the traffic management body, collect members, and then start issuing orders: I can call myself a football expert, and perhaps even find some credulous acolytes who will accept my decisions, but that doesn't mean I get to pick the Scotland team. There has to be some link between my authority and how I came to have that authority; or, to put it another way, the obligation (demand, directive, etc.) has to be comprehensible in light of the institution.

Perhaps, if we assume that my reckonings about the best XI available to Scotland are correct, I have some limited epistemic authority. But this gives me no institutional authority, and no remit to call up players: unlike the national team's management, nobody wants (or expects) me to pick the team. When the manager issues a summons, they *do* generate a (weak) obligation, and do it in virtue of their institutional role - a player who isn't called up ought not to turn up anyway on the basis that they, even correctly, think they'd be better than one of the actually selected players. The flipside of this is that the manager is

¹¹ Here's an obvious case: there are prudential reasons for me to drive within the speed limit, mostly to do with my safety. If that is the *only* sort of reason for me to do so, then there is no principle stopping me from mounting a crowded pavement if I'm being chased by an (automotive) axe-murderer.

assumed to be making their picks according to reasons consistent with their holding the position. If this is not the case - if, for example, some player is repeatedly capped because of their personal relationship with a coach - then the manager's authority is accordingly called into question. In short, then, the kind of institutional obligations that can be generated by the Scottish Football Association (in this case) are both justified and constrained by the purpose of the institution.

This is a fairly familiar account of one aspect of political obligation (see for example Hanrahan & Antony op. cit., Klosko 2005, Mokrosinska 2012, Raz 1986), and in order to make progress in understanding whether some institution actually *is* authoritative according to the conceptual description given here we will have to know how the institution is set up, whether its procedures are consistent with its purpose, and so on. This task can be left until later. For now, we should say that a political authority may generate obligations in virtue of its institutional structure if

i) That structure is consistent with its purpose, however this is cashed out; it would be difficult to see why anyone should pay attention to national call-ups if the selection process was handled by a random assortment of folk uninterested in football.¹²

ii) The obligation would not apply (or be comprehensible) were it not for some quality or qualities of the institution - again, the SFA composed of passers-by seems apt to fail here.

iii) The obligation's issue meets whatever internal or procedural requirements delimit the institution's purpose. Return again to the football example. If we had a selection committee composed of seasoned ex-pros, scouts and coaches, all aiming to pick the best team possible, and one member performed some *fait accompli* with regards to announcing the squad, then the announcement should not be regarded as authoritative or even being generated by the SFA. Despite satisfying the first two criteria in virtue of the SFA's structure being (by

¹² Although not impossible by any means: perhaps we are more concerned that the team should be truly representative rather than (notionally) determined by skill levels, in which case a Buggins' turn system of selection would seem much less bemusing.

stipulation) consistent with its purpose, and the squad being comprehensible in light of the institution's purpose, the rogue announcement is not an institutional obligation because it has not passed the relevant procedures, whatever they might be.

This subsection has shown that the conceptual requirement for putative authorities to generate obligations by virtue of their institutional structure implies further justification and constraint requirements. For political authorities, we must be able to understand their structure in light of their function to promote autonomy; their ability to issue obligations must cohere with that structure and function; and the directives that they actually issue must pass whatever procedural constraints are placed on an autonomy-promoting institution.

2. *Institutionally-Generated Obligations*

The desideratum, to recapitulate, is that obligations generated by the authority get their force from something about the institution itself: they are not just institutional assertions of free-standing moral requirements. To use a well-worn example: the reason that I ought not (*ceteris paribus*) to murder people is because it is wrong, not because there are laws against it. When (e.g.) the British state tells me not to murder, I *comply* rather than *obey*: I act in accordance with the law, but not because of the law. That is to say, the authoritative nature of the command "do not murder" has nothing to do with the issuer, and everything to do with the content.

If there is to be an institutional *political* authority, however, it seems that it should have something aside from (although perhaps not 'over and above') purely derivative moral force of this sort. The reason for this is simple: much of the work of co-ordinating groups of people will involve resolving blameless disagreement rather than Manichean conflict. There is, laying aside conceptions of justice and retribution for a moment, no manifest need for any body to adjudicate that murder is wrong; but there is a manifest need for some method of arbitrating between competing proposals (over, for example, some shared space). Unless we are to adopt a kind of optimistically but narrowly technocratic

or perfectionist model whereby there is one and only one best use for any space or resource - and a group of people who can reliably identify it - then any decision regarding these proposals will not be able to coat-tail on the moral status of the proposals themselves.

But we still, on pain of falling into the individualist trap discussed earlier, want the eventual decision to bind us in a stronger way than the advice of a friendly expert. It should not, under normal circumstances, be morally permissible for the relevant institution to rule “out of *x* or *y*, *x*” and yet for me to continue *y*-ing.

The worry, to briefly sketch it here, is that even obligations that are in *form* the result of institutional declarations - such as traffic laws in the UK - may get their *force* from free-standing moral reasons. There is no independent moral reason for anyone to drive on the left (or right), the thought goes, but given the convention of driving on the left, to drive on the right is to inflict inconvenience, or more likely risk, on others in a way that is ruled out by everyday moral rules.

If it turns out that this is true of any purportedly institutional obligation, then I will have to give up the strong form of the second desideratum. Even in that eventuality, however, note that we will still have a reason to do *that* rather than *this* on the basis of institutional say-so - albeit that the reason will be an indirect one resembling “the institution says do *that* (and there are moral reasons to endorse the institution’s instructions)”. In this case, whether or not an institutional directive was morally binding would be more dependent on whether following the directive was the best (or best available) way of pursuing other goals.

Perhaps, then, there is no clear distinction between how an institution might go about generating obligations, and how those obligations come to be morally binding; a thought which is supported by the following example.

Imagine that we are members of some friendly association that holds a monthly dinner in one or other of its members’ homes. Everybody puts in a certain fixed sum to cover the dinner, which always falls on the first Sunday of a month. Now,

the obligation to pay in such-and-such an amount by this particular date is institutional in form: not only are the specifics determined by the institution, but the obligation would not even exist absent the council of diners (or whatever we wish to call them). Is there any moral dimension to the obligation? Plausibly so: whether we argue from fairness, from benefits received, or any of the possible contenders, it seems that there is at least a *prima facie* wrong I commit (whether or not I have explicitly consented) if I take the money and don't cook, or turn up at other dinners without paying my subs.¹³ If I wasn't a member, then the obligation wouldn't apply at all; if I kept being the one selected to cook then, plausibly, the obligation weakens or again falls away. Crucially, it does so not because of any free-standing moral reasons or obligations - that is, it's not obviously wrong for others to expect me to cook once a month - but because the institution, such as it is, is supposed to produce directives consistent with the presumed aim of giving everyone a turn to cook and turns to simply enjoy the food.

The institution's being able to produce binding obligations, in other words, is closely correlated with the institution's making appropriate demands. In the dinner case, only those directives relevant to performing the institution's function of monthly meals will be apt for authority; in the traffic case, the candidates for authoritative status will be all and only directives relevant to affirming a convention regarding which side of the road to drive on. More generally, whether or not some putative obligation is indeed binding will often be a substantive matter of whether the obligation will (or is likely to) serve the institution's purpose, and whether the decision-making procedure which produced the directive is appropriate for that purpose.

3. *Coherence with Autonomy*

The last requirement is again a simple one: if we create an account of political authority which is plainly inconsistent with the account of autonomy I have given - "an institution is authoritative iff it does god's work", and so on - then one of the two has to go. Either the kind of institution formed would fail in promoting

¹³ I discuss Klosko's fairness argument (which coincidentally also involves a gathering of gourmands) and Nozick's objection to obligations from benefits received, later in the chapter.

and protecting autonomy, or the conception of autonomy I have given is vitiated by requiring a hyper-voluntarist view of social institutions. This is an external failure, in that neither concept need cause any problems by itself but only generates inconsistency when placed with the other (we might also think of this as a higher-order requirement).

This requirement is the most difficult to describe in abstract rather than substantive terms. However, by considering the desiderata, we should at least be able to see what *role* the requirement plays (instead of seeing how exactly the requirement may be met).

Perhaps, in my schema, it serves as a high-level conceptual constraint; we cannot know whether some putatively autonomy-promoting institution *does* generate obligations of the correct sort without substantive investigation, but we can know that any obligations it issues which are inconsistent with autonomy flatly fail to be authoritative. There does, however, seem to be one aspect of authoritative directives which only this requirement heeds: namely, their implementation. It is clearly possible for some directive to pass through 1 and 2 while failing 3 by dint of what would be required to enforce it: if we take the inoffensive “drive on the left” obligation, it would still be possible for such a directive to be non-authoritative if the punishment for breaking it were something like familicide.

One further thing to be said at this point is that we can understand “consistency with autonomy” in at least two ways. The first is that the directive straightforwardly increases or does not reduce the number (or efficacy, or range) of autonomy-constituting relations that agents stand in. The second is that it makes some relations unavailable in exchange for creating other, more numerous (more effective, wider-ranging) autonomy-constituting relations.

The first interpretation requires that, in essence, every directive must leave untouched all autonomy-constitutive relations it does not improve in order to be authoritative. This is exceptionally demanding, but the more serious problem comes when we consider that such a move leaves us engaging in a kind of false consciousness defence when apparently valuable relations are changed or

removed. That is to say, if there is some decision which (for example) requires that 20% of one's garden space is used for vegetables rather than decorative flowers, then it seems reasonable to say that something of value *has* been lost for the person who can no longer solely grow their preferred roses. The decision may still be legitimate and entirely justifiable by reference to autonomy, but it still seems that the rose-grower experiences some disvalue. If we are committed to the first interpretation, then we must be committed to the claim that the growing of roses just ceased to be valuable self-direction when the directive was issued (or that it was never valuable in the first place); to believe otherwise is a form of false consciousness. Akin to the Rousseauian line on dissenters from the General Will, or the orthodox Marxist characterisation of anti-communist proletarians, the agent has mistakenly or as a result of manipulation come to hold a belief which is against their own interests.¹⁴ In this case, the gardener is simply wrong about rose-growing being a valuable part of their life. This doesn't seem like an absolutely insurmountable problem, despite its initial unpalatability (especially since I'm committed to something awfully like a false consciousness response with regards to, say, folk who take autonomy to be instantiated simply by unrestricted choice), but it is certainly a theoretical burden.

On the second interpretation, we need take on no such burden: sometimes institutional directives will, in fact, deprive us of autonomy-constitutive relations. If the directive creates or is likely to create more such relations, then the command is still authoritative. We have committed to maximising equal autonomy, and this may sometimes involve curtailing unobjectionable (perhaps even valuable) self-direction; it is a trade-off, but so long as the balance tips towards more rather than less autonomy, it is a trade we are obliged to accept. This second account of how to reconcile autonomy-justified authority with limitations of self-direction is, again, fairly common in the literature: our familiar stalking-horse Raz clearly holds such a view (op. cit., pp410-412), and Oshana's observation that "if we distinguish global autonomy from local autonomy [or as I would have it, programmatic from episodic], it may be that

¹⁴ See Rousseau (1968) for the former and, among others, Gramsci (1971) and Eyerman (1981) for overviews of the latter.

global autonomy is one of those values for which local autonomy or personal liberty can be sacrificed” (2006, p120) strongly implies it.

This, I think, is about the limit of what we can say about the third requirement without getting into the nitty-gritty of what sort of things actually are or are not consistent with autonomy. Although there will obviously be some proposals that are obviously conceptually incompatible with maximal equal autonomy - fascist directives, racist or sexist proposals, etc. - it seems likely that many or most disputes about whether or not a given policy will actually be consistent with autonomy will be resolved on the basis of empirical and contingent factors. That is, whether a proposal like “everyone should spend an hour a week x-ing” is actually consistent with autonomy will be dependent not only on what x is, but on the context the proposal is made in and the penalties (if any) for non-compliance. If x is something like “planting potatoes”, it might be consistent with autonomy in a society short on food, or indeed a society which preferred localised food production where possible. But it could equally be inconsistent if the punishment for disobedience was death, or if many members of the society lived in places where growing potatoes was impossible: since, in the latter case, it would effectively be an injunction to do something which limited folks’ self-direction in a way which - as it happens - instantiated and promoted no autonomy-constituting relations, and hence would not be consistent with maximal equal autonomy. The question of what broad sorts of substantive policies are likely to meet this third requirement is one that I return to in the next chapter.

3. Some Competing Conceptions

We now have the desiderata which a concept of authority must possess in order to play the role required of it. Before putting forward my account in the next chapter, considering other apparently promising conceptions of political authority will allow us to make headway on what is required for an account to fulfil the three requirements for a concept of authority.

In this section I examine and criticise two particular conceptions of political authority: Raz's Normal Justification Thesis (NJT hereafter) and Klosko's argument for political obligation rooted in fairness. Neither account, I argue, can satisfy all three conditions for a political authority's being legitimate, although both contain useful and broadly plausible claims about authority. In particular, Raz's notion of dependence - that the reasons we have to obey an authority should be in some way dependent on reasons we are already committed to taking as reasons for action - and Klosko's argument from fairness both speak to the autonomy-promoting function of authority that I have argued for thus far.

3.1 Raz and the NJT

One immediately helpful aspect of Raz's account is that he adopts a "normative-explanatory account of the core notion of authority" (op. cit., p65) which is roughly the same as what I have called the functional analysis: he takes the concept of authority to be one which we shape and deploy for a particular purpose and in particular contexts, rather than having an antecedent and fixed meaning which we access through careful conceptual analysis.

The first step of Raz's argument involves identifying a crucial feature of authority: if we take a putative authority's command as merely advice or suggestion, then the command is not authoritative. If institution A really is authoritative, then their telling me to do or not do something is a reason to act (refrain from acting) that simply over-rides any judgements I might make about the action. In Raz's terminology, legitimate authorities give us pre-emptory or exclusionary reasons, reasons which should stop our reflection about whether or not to act at the point where the authority issues its command. An institutional command "is not to be added to all other relevant reasons when assessing what to do, but should exclude and take the place of some of them" (p46).

If authoritative commands are - as they plausibly seem to be - pre-emptive in this sense, and if Raz can show how such commands can come to exist, then we have two out of three features required by the concept of authority. A reason that replaces or over-rides other reasons for action, and does so specifically because it is provided by (or perhaps is constituted by) an authoritative

command, is clearly an institutionally-generated reason. Not only would the reason not exist absent the institution, but the reasons that should motivate us to act would be entirely different absent the institution.¹⁵ Equally clearly, a directive that pre-empts other reasons is at least of the form of a morally-binding obligation. The mere fact of some reason for action being exclusionary does not automatically make it a moral obligation. Consider the authority of a referee in sport: we shouldn't retreat ten yards for a free-kick iff we think the referee's decision is correct, but neither would we think that under normal circumstances disobeying the ref was a moral failure. Nonetheless, it does seem that if we take there to be some underlying moral reason to act in accordance with the institution's directives then we must also take the exclusionary reasons it provides to have moral force.

Leaving aside for a moment the question of consistency with autonomy, the task for Raz would thus be to give us an argument to think that political authorities can provide exclusionary reasons of this kind. To do this, he offers two theses: the Dependence Thesis and the NJT. Since the latter is reliant on the former, I will deal with the Dependence Thesis first.

The Dependence Thesis, as described by Raz, claims that:

All authoritative directives should be based on reasons which already independently apply to the subjects of the directives and are relevant to their action in the circumstances covered by the directive (p47)

In other words, an authority cannot just drop out of the firmament and start ordering me about because I happen to fall within its claimed territory; I have reason to obey the authority (or the institution is genuinely authoritative) only if I am already committed to some principle or course of action which the authority and its directives are concerned with. To give an example, directives about which side of the road we drive on may be authoritative because they are based on reasons - ease of transport, public safety, etc. - which already apply to

¹⁵ To be clear: if an institution says "Don't murder" authoritatively, then it isn't the case that my previous reasons not to murder just cease to exist; rather, I never get to the stage where it is appropriate to reflect on them. It is in this sense that I take Raz's account to mean that the reasons that should motivate us are different under the institution, a sense supported by his comment that "the difference is not in the presence of an additional reason for action, but in the existence of a pre-emptive reason" (op. cit., p60).

us or, in my terminology, because we are committed to a safe society (for example). Given that these reasons already apply to me, and given that conventions about road usage are obviously relevant to public safety considerations, a directive which says “drive on the left” satisfies the Dependence Thesis.

The Dependence Thesis goes some way to explaining why, for example, some church’s imagined claim “parking on double-yellow lines displeases God” seems to fail at being an authoritative directive. Firstly, there is no independent reason for me to care about what pleases or displeases god; secondly (and relatedly), whether or not something pleases god is not relevant to my deciding whether or not I should risk parking on double-yellows. The directive gives me no action-guiding principle because it is not based on any reasons that apply to me. Note that, if we were to assume that pleasing god *was* a reason which independently applied to us - in the same way that moral reasons are taken to - then the Dependence Thesis would suggest that such a directive could be authoritative.

It is worth being clear that the Dependence Thesis gives us a necessary but not sufficient condition for authority. We can easily imagine a situation where somebody gives me relevant advice or suggestions without this constituting an authoritative command. If my friend tells me that such-and-such chocolate bar is produced by a more than usually evil company (and thus I shouldn’t buy it) as I consider my options in the newsagent, then they are offering me a relevant consideration on the basis of reasons that I am committed to, in this case reasons involving not supporting especially obnoxious businesses. But this manifestly does not put my friend in the position of being a political authority over me. If they have any authority, it is the moral authority attached to making correct moral claims, but we have already established that such authority is not the point of contention here.

Rather, if a putative authority issues a directive based on reasons that don’t apply to the relevant agents, and/or if the directive includes no relevant action-guiding imperatives, then we can disregard it even if the authority is set up in such a fashion as to generally provide appropriate directives.

Two worries arise here, the first exegetical and the second conceptual. The exegetical worry is that I have simply misrepresented Raz: he explicitly denies that authorities have to get it right on every occasion, that “authoritative determinations are binding only if they correctly reflect the reasons on which they depend” (p48). How can this be squared with my claim above about when we can write off directives as non-authoritative?

Curiously, explaining the exegetical worry goes some way to addressing the conceptual worry, which is about how we can know when authorities fail to issue authoritative directives and how our exercise of such knowledge coheres with the Pre-emption Thesis. Although Raz indeed denies that an authority’s being mistaken necessarily makes it illegitimate, he argues that some kinds of mistakes have just this effect. This invites the conceptual worry, which he puts in the following way: “[if] the directives [of a putative authority] are binding only if they do not deviate much from right reason and as we should act on them only if they are binding, we always have to go back to fundamentals” (pp61-62). In short, if we cannot know whether or not an institution has made a mistake of the sort that justifies disobedience without considering the issue on its own merits, then the Pre-emption Thesis no longer applies, and authorities cannot play the role that we want them to.

Raz’s response is to attempt to draw a distinction between “great” and “clear” mistakes, with the parallel of bungling a long process of addition. In the first instance, we might be out by many thousands without the mistake being detectable except by painstaking recalculation - a great, but by hypothesis, not clear mistake. A clear mistake, conversely, would be something like an answer which included a decimal point despite only integers being added: anyone with basic mathematical competence would instantly see that something was wrong without having to run the whole calculation. We can, Raz implies (*ibid*), only challenge authoritative directives when they have made a clear mistake, and unless we think that the majority of mistakes are clear (or that it is difficult to distinguish between a clear and great mistake) then we need not worry that the Pre-emption Thesis is ineffectual in practical terms.

We may or may not think that this is a satisfactory distinction, and even if it is we'll need a fairly detailed picture of what a *clear* mistake would be for the autonomy-promoting authority that my thesis seeks to define; but that is a matter for the next subsection.

So far, then, we have established two of Raz's key theses. The Pre-emption Thesis says that legitimate authorities give us pre-emptory (or exclusionary) reasons for action, reasons which supersede our normal processes of evaluating what to do in particular situations. The Dependence Thesis stipulates that these exclusionary reasons must ultimately be taken to apply to agents because they are based on independent reasons which already apply to those agents - that is, authoritative directives should be dependent on principles (values, reasons/reasons for action, etc.) which are relevant to me, and should provide action-guiding advice consistent with my following those principles in this situation.

Taken together (although it should be noted that this is not how Raz's dialectic proceeds), the Dependence and Pre-emption Theses suggest a more general claim about authority, the well-known Normal Justification Thesis. This states that, to quote:

the normal way to establish that a person has authority over another person involves showing that the alleged subject is likely better to comply with reasons which apply to him (other than the alleged authoritative directives) if he accepts the directives of the alleged authority as authoritatively binding and tries to follow them, rather than by trying to follow the reasons which apply to him directly (p54)

The NJT implies both the Dependence and Pre-emption Theses, although only weakly in the latter case. If it is in general true that authorities can only demand that I act for (authority-independent) reasons which already apply to me, and there are plausible motivations for thinking that some of these reasons are best served by my not analysing each putatively authoritative directive, then we can (as Raz does) extend this claim to the stronger pre-emption claim that if I am committed to some value, then being told what to do by some body which reliably serves that value should simply shortcut my usual evaluations. Again, a

football analogy might help: if I am a centre-back committed to helping my team win, and my goalkeeper is generally better-placed to arrange the central defenders in a way likely to prevent the opposition scoring, then the goalie's shouting at me to mark so-and-so or cover such-and-such part of the box should be enough to make me mark or cover the relevant targets. I act in some sense inconsistently with my stated commitments if I deliberate on the goalie's command rather than following it; so too, says the Pre-emption Thesis, with legitimate political authority.

So much for exegesis. To put this in terms of how the NJT matches up to my proposed desiderata for political authority, the three questions to ask of it are:

1. Does it explain how institutional directives are (or may be) morally binding?
2. Does it explain how these directives get their bindingness from the institution itself (rather than being independently binding directives that the institution simply parrots)?
3. Is it consistent with the relational conception of autonomy that I have defended earlier in the thesis?

In the next subsection, I argue that the answers are, respectively, yes, yes, and no.

3.2 The NJT and the Three Desiderata

We have already seen that authorities which are legitimate according to the NJT should be able to answer the first two questions above in the affirmative: the directives of legitimate authorities are morally binding insofar as they command on the ultimate basis of (normative) reasons which we ourselves are or should be motivated by. For example: Institution A, a traffic management committee, tells me not to park on double-yellow lines, and if I have good reasons to believe that institution's directives generally promote or defend autonomy-constituting relations then I am morally bound to obey.

The requirement that this obligation be dependent on the institution (rather than the institution happening to issue orders which run parallel to independent moral requirements) is also met by an authority which satisfies the legitimacy conditions of the NJT. In addition to the thought that directives of convention definitely *do* take their particular form from the institution, the reason that we have to obey an institution is not *solely* reducible to (though it is ultimately dependent on) the reasons to promote autonomy that we already have. There will be background requirements of constitution, process and the like which mean that we have not just directives the form of which are institution dependent, but directives which would not, *even if identical in form*, be authoritative when issued by another institution. That is to say, even if Institution B (a religious body to be contrasted with A) issues a command which as it happens would be the best way to promote autonomy, it is not an authoritative directive; in Razian terms, we have no reason to think that our reasons are best served by obeying B, nor should we take B's pronouncements to pre-empt our own evaluation of the matter.

So far, so good - the ways in which institutions can promote autonomy, or how we come to have reasons to believe that they do, will have to be cashed out and are likely to be very different from the Razian statist view, but as a conception of authority which is consistent with the demands of the functional analysis the NJT seems promising.

However, the question of whether the proposed conception of authority can accommodate (or be accommodated by) my account of autonomy prevents us from simply adopting the NJT wholesale and giving it an anarchist hat, for it does not seem to me that maximal equal autonomy is consistent with the really very rigid and hierarchical implications of the Pre-emption Thesis as Raz gives it.

On the Razian model, it seems as if the only instance of 'clear' mistakes - those mistakes which *do* allow us to reflect on and reject the purported authority of an institutional directive - are those which he terms jurisdictional mistakes (p62). That is, it's only when the authority tries to direct our behaviour in a realm over which it has no proper remit that we can reject the command on the basis of our own judgement (or, indeed, apply ourselves to any serious scrutiny of the

command in the first place). Leaving aside for a moment the question of whether it is really as clear and easy to detect jurisdictional mistakes as Raz suggests, the thesis carries the implication that our own reasons would be best served by following the directives of an institution which consistently failed - indeed, failed greatly - to reflect them, just so long as that institution refrained from making any jurisdictional mistakes.

Recall that for an agent to be autonomous, on my conception, it is necessary (although not sufficient) for her to be non-dominated; which is to say that she must stand in relations such that any arbitrary or unjust attempt to interfere with how her life goes will be defended against, and that this defence must be grounded in recognition respect of her as an agent. One of the features of recognition respect is at least *prima facie* willingness to grant that the agent is an appropriate source of values and can make moral judgements. We may, of course, over-rule the agent if they are acting immorally, and perhaps in certain circumstances if they are acting unknowingly against their own interests, but the presumption of agential primacy is a fairly strong one. In short, we tend to think that the only time it is permissible to interfere with or coerce the agent is when they are acting inconsistently with the equal autonomy of others.

This on its own is not an obvious defeater of the NJT. If we take the plausible stipulation that following an NJT-legitimated authority would tend to promote maximal equal autonomy, then someone's rejecting an institutional directive would be grounds for *some* measure of interference, so (at the risk of repeating myself) having one's self-direction restricted is *not* necessarily inconsistent with autonomy.

The problem is instead that having one's self-direction interfered with, or standing in relations such that it may be interfered with, by an authority acting on neither moral nor 'dependent' reasons (that is, in pursuit of reasons which apply to the agent) looks very much like being dominated. Indeed, on Raz's view even to try and evaluate some directive from an NJT-legitimated authority, let alone reject it, under one's own steam is to fail in a duty of obedience. Let us take an example where we specify "the factors about which the authority was wrong, and which are not jurisdictional factors" (p62), and where the Pre-

emption Thesis would tell us that we are nonetheless obliged to obey without engaging in evaluation.

The case is one where the authority is, on the whole, generally reliable as an autonomy-promoting institution; by and large, we are likely to better comply with reasons which apply to us by complying with the directives of the institution. In this instance, however, the institution issues a greatly mistaken directive that says some proportion of public land must be turned over to growing quinoa (instead of what it was growing before). The reasoning is that quinoa will be more effective at feeding the citizens of this region than the alternative, thus freeing up labour time for other tasks or for relaxation - and so, the thought goes, promoting a state of affairs which is likely to bring about greater equal autonomy. Let us add, also, that the institution is either wholly concerned with co-ordinating regional crop rotation, or that this sort of thing properly falls under its aegis.

The reasoning is plausible (we stipulate), but as it happens wrong: replacing the current crop with quinoa would be far *less* efficient, resulting either in extra work or, in the worst case, widespread food shortages. Although the reasoning is wrong, it is not wrong in any immediately clear way - there is no stage of the calculations where the committee have thrown in "number of days since my birthday" or anything similarly arbitrary as a variable - and, as we have specified, the committee makes no jurisdictional mistake. In Razian terms, it is a great but not a clear error, and so it seems that we are obliged to obey the directive. Even those rascals who *do* apply their own judgement, and realise that the calculations are disastrously wrong, are under the same obligation.

But look at the structure of what is happening here: even though the agent who refuses to farm quinoa is not doing anything contrary to maximal equal autonomy, and even though they have good (indeed, correct) reasons to justify or account for their behaviour, the Razian model suggests that they be coerced in the name of autonomy. There is no appeal to one's own reasons; the agent cannot cop to a *prima facie* breach of duty but point out that their actions are ultimately justified, because they are not supposed to judge for themselves ("clear" mistakes aside) when the *prima facie* obligation ceases to hold. In other

words, the explanation for their behaviour, if they were to follow the directive, would have no room for their own reasons - they stand in relations such that their lives can be interfered with without reference to the agent's values. Again, this looks like domination, and that means that the Pre-emption Thesis fails the requirement that conceptions of authority be consistent with the foregoing account of autonomy.

There is an obvious Razian objection to field here before moving to suggest a weaker variant of the Pre-emption Thesis. The objection runs as follows: to claim that coercing the dissident farmer in the quinoa case is not defensible by reference to autonomy is to beg the question against the NJT, or at the very least to overlook a crucial fact about it. Raz is clear that the reason we are obliged to obey even those directives which are greatly mistaken is because without the Pre-emption Thesis, no institution can perform the co-ordinating role required to fulfil part or all of its function.¹⁶ Thus, the motivation behind over-ruling the recalcitrant pastoralist *is* an autonomy-minded one - if everyone adopts an attitude of non-compliance towards directives they disagree with, the institution will be ineffective in its goal of promoting and defending maximal equal autonomy - and, consequently, one that is based in the agent's own reasons.

It would be disingenuous to suggest that the problem here is the possibility of being over-ruled against one's firmly-held (in good faith, after rational reflection, and so on) principles. As I have made clear elsewhere in the thesis, my account will forbid the earnest, authentic right-libertarian's attempt to remove themselves from collective obligations, and will even do so on the basis that what they think about autonomy is substantively incorrect. My objection to Raz cannot, therefore, be an anti-perfectionist one. Equally, the Razian thought that an authority which has no pre-emptive force is no authority whatsoever seems largely correct - so I cannot object to the mere over-ruling either.

¹⁶ I have argued that resolving 'mere' co-ordination problems is not the only purpose of authorities, but it will be simplest if we take co-ordination to be the sole desideratum for the sake of this example. We might in any case think that sometimes the resolution of co-ordination problems will be (as it happens) the sole requirement for some policy's having an autonomy-promoting impact.

Rather, the problem is that the Pre-emption Thesis as given reduces the justification for political obligation down to one we might be given by a particularly unimaginative parent; we ought to obey the law *because it's the law*, and we ought to obey a putatively authoritative institution because it's a putatively authoritative institution. Just as "because it's the law" is a directive which, insofar as it has any power whatsoever, is reduced to empty authoritarianism the moment that law obliges us to do something immoral (or refrain from acting morally), the notion that we should obey an institution because of autonomy, even when the institution is *knowably* acting in a fashion which is hostile to autonomy, renders the whole business of searching for plausible justification moot.¹⁷ To draw the point out, the quinoa example provides at least two reasons to think that the Pre-emption Thesis as given won't work for my purposes.

First, the agent's values involve (by hypothesis) the authority *actually* serving the goal of maximal equal autonomy, and when the institution fails to do this, it fails to reflect the agent's reasons. In other words, it is not enough for the institution to follow the right kinds of procedures; it must also reach the right outcome in order to be authoritative. Admittedly, in many cases, following the correct procedures will give the right outcome - for example, if we ask the appropriate people about how best to use some patch of land for recreation, record their responses accurately, and act in accordance with those responses, that strikes me as the outcome likely to produce maximal equal autonomy. But this is not the case in the example under consideration, and it is also not the case here that compliance will end up being more autonomy-conducive than non-compliance.

Second, the problem is not just that the authority gets the wrong outcome; indeed, the more pressing issue is to do with the relationship between agent and

¹⁷ Someone sympathetic to Raz might respond, as Ben Colburn has suggested in discussion, by modifying the boundaries of what it is to be a clear mistake so that any policy incompatible with maximal equal autonomy is guilty of making a clear mistake (just as a jurisdictionally-illegitimate policy would be). This response would certainly enable them to evade my charge as given above, but does not make the NJT overall more appealing for my purposes; since, as I have suggested, our right to disobey cannot be limited solely to instances of clear mistakes. If an authority makes a great mistake, we should be able to challenge it on the basis that it is no longer serving the goal of maximal equal autonomy, not just on the basis that it is *clearly* no longer serving that goal.

authority on Raz's model. We have stipulated that the institution is wrong, and that the agent *knows* they are wrong - in other words, that the authority is not best serving autonomy by its directive. But this is not, for Raz, enough to justify the agent's refusing to comply, or to undermine the institution's authority. This means that not only does the authority fail to reflect our values with regards to the specific quinoa directive, it also fails to reflect our values with regards to how authorities should be directed - it gets the wrong result *and* we cannot (or ought not) challenge its decision. The institution interferes with our lives in a way that does not serve our values, and we're required to obey it anyway; this looks to me exactly like domination.

We'd be as well, if we buy the Pre-emption Thesis as-is, taking a Hobbesian line that *any* political authority is an improvement (autonomy-wise) over a world without authority; it would at least have the virtue of honesty.

Even if we think that such a line is correct, it is clearly inconsistent with the anarchist motivations of my thesis. I have already indicated that one fail-state for the thesis would be the conclusion that a Nozickian or Stirnerian individualism is entailed by philosophical anarchism; my ship is sunk if it transpires that a rejection of state authority must also involve an appeal to the primacy of individual autarchy and hence that there can be no legitimate political obligations which we have not explicitly consented to. But the conclusion that we ought to obey institutions (and that institutions may legitimately coerce us) regardless of considerations of autonomy is surely another equally unappealing fail-state. It's difficult to imagine a more comprehensive betrayal of the anarchist intuition than the one which leads to "my committee, right or wrong", and an account of authority which forbids us from critically examining the directives of a putative authority looks prone to embracing exactly this kind of dogmatism.

This doesn't, though, exhaust the Razian's options in responding to my criticism. In the same way as the previous objection suggested that I had overlooked the authority's purpose - that, in effect, I was assuming that it was coercing people just for the sheer hell of it and then charging it with unjustified coercion when it was in fact engaging in coercion in order to promote autonomy - this objection

would accuse me of mischaracterising our relationship with the putatively authoritative institution. I have already characterised the institution as one which has previously been fairly reliable (and known to be reliable) in terms of introducing autonomy-promoting policies, and Raz has already claimed that “trust in the authority is trust that the authority is likely to discharge its duties properly” (p55). In light of this, we must surely then think that there are reasons to trust the authority. It is not as if this institution suddenly appears, delivers a (mistaken) demand about farming and then expects obedience. We have, by stipulation, lots of historical reasons to think that following this institution’s policies *is* in general good for maximal equal autonomy. In other words, the interlocutor can object, we have a vast array of reasons to obey the putative authority, even if we *know* that on this occasion its policy is incorrect, and talk of “obeying the law because it’s the law” is (albeit not quite incorrect) by no means the whole picture. We ought to obey *this* law because it has been laid down by *this* institution, and said institution is on the whole an autonomy-promoter: it may, indeed, be non-ideal, but it’s not a blind-faith commitment.

There is something to this, but it will not do the job that the critic wants. If I have good reasons to trust an authority, and no (or weak) reasons to doubt it in this particular case, then it does seem as if I ought to obey a policy even if I am sceptical of it. But in the quinoa case, to re-flog the horse, we have an indefeasible reason to think that the institution will *not* discharge its duties properly - namely, that we know for a fact that its policy is mistaken and will not have the effect of promoting autonomy - and so the benefit of the doubt that the institution has accrued over the years simply doesn’t come into play. If I correctly inform you ninety-nine times that a bay window is open and you can run through unimpeded, then you have good reasons to take it on trust when I tell you the same thing on the hundredth occasion. But those reasons are defeasible: if you can see your reflection as you approach what I assert is empty space, then you’d have to be a blessed innocent or a bloody fool to keep running, and this is ultimately what the Razian version of the Pre-emption Thesis seems to me to demand.¹⁸

¹⁸ It will do no good to claim that we should still obey the policy because (say) widespread dissent will undermine the institution: if the institution is generally on the ball, then the

This, of course, leaves me with the task of explaining how pre-emption ‘fits in’; a task which must largely be left until the next subsection. We can, however, note at this point that something like the Pre-emption Thesis seems to work as a constraint on baseless or bad faith rejections: that is, unless an agent can give some plausible account of *why* they take a putatively authoritative directive to be wrong, they must take themselves to be pre-empted. Intuitively, there is a difference between “I think your command is mistaken because *x*” and “I don’t *wanna!*”, and it seems that, at the very least, the latter rejection fails to give any meaningful reason for rejecting authoritative pre-emption. If we cannot give some account of why rejecting a particular directive seems likely to result in our better complying with our own reasons then it is difficult to see what reason we have for disobeying the directive.

3.3 Authority and Obligations from Fairness

I have argued in the previous subsection that the NJT cannot play the role required of political authority because the Pre-emption Thesis is incompatible with the relational view of autonomy (and so the NJT is inconsistent with an autonomy-motivated conception of authority). Ultimately, the Pre-emption Thesis tells us that we should obey putatively authoritative institutions just because of their putative authority, even if we have excellent reasons to think that their authority is absent or critically undermined in some cases. Although, as suggested, the Pre-emption Thesis might account for why we dismiss bad-faith refusals of institutional directives - if you have no reason to think the authority is wrong, it’s not clear why you should take it to be wrong - it does *not* give a motivation for why we should obey such institutions. That is, while we can point out that mere suspicion or disinclination is a bad objection to some directive, we cannot motivate the recalcitrant agent by appealing to autonomy. By stipulation, after all, they’re not convinced that the directive can be justified in such a fashion, and it’s both easy and plausible to imagine a less apparently-truculent sceptic than the baseless foot-stamper.

occasional spate of justified refusal ought not to destroy it, and if it’s persistently making mistakes then - as suggested in responding to the previous objection - suggesting that we ought not to undermine it doesn’t seem attractive unless we want to call the game a bogey and join Hobbes in his bunker.

Take some directive which will impose a noticeable, albeit mild, burden on all members of the community, and which has met whatever procedural standards are required of it. There would likely still be those who are unconvinced: perhaps there has been a high-profile mistake recently, or there is a great deal of division in public opinion about whether the directive is truly optimal, such that our hypothetical agent can be described as blamelessly ignorant or indecisive. In such cases, we might want to appeal to a principle which demands compliance because, recalling Antony and Hanrahan's "Gigli problem", *some* decision has to be made and non-compliance with the decision will result in the decision being effectively unmade. A return to the initial deadlock would thus loom, and the benefits of an authoritative institution would be lost (at least in this case). When everyone else has, or a majority of others have, willingly accepted the burden in order to maintain this public good, it seems unjust that a resistant minority, even a sincere one, can reject the obligation. We would, in other words, be arguing for obligation not on the basis of asserted authority, but on grounds of fairness.

There is of course a tradition of liberal political philosophers making arguments from fairness or fair play, perhaps most importantly Hart (1955), Klosko (1987, 1994) and, although he does not think that principles of fairness alone can generate political obligations, Rawls (1964, 1971). In this subsection, I focus mainly on Klosko's development of the principle.¹⁹

First, "suppose there is a mutually beneficial and just scheme of social cooperation, and that the advantages it yields can only be obtained if everyone, or nearly everyone, cooperates" (Rawls 1964, p9). The principle of fairness in such a scheme is characterised by Klosko's claim that for some agent A, "if others are willing to co-operate and the benefits in question are provided, A is also obligated to co-operate" (1987, p354). This seems intuitively plausible: I ought not to sit about while others pick apples and then gorge myself on the results, to give a very basic example.

An immediately obvious worry affects this formulation, however. If I have no desire to eat apples, or at least no desire strong enough to motivate me to

¹⁹ For a useful overview of Hart's and Rawls' arguments from fairness, see Simmons (1979).

action, then it's not obvious why I should take part in the tree-climbing, even if I consequently eat an apple (and thus receive the benefits). I would, after all, have been just as happy without the apple; it is not that I've waited to pounce and begone with my ill-gotten fruity gains, but that I am so minimally interested in eating apples that I will only pursue the goal if it involves next to no effort. This is not something that we generally regard as untoward behaviour in everyday life: variations on "if you're going to the shop, I'll have an X, but I can't be bothered getting one myself" are both common and unobjectionable, so it can't simply be that receipt of benefits *simpliciter* makes one obligated under a principle of fairness (I do not engage with arguments about receipt versus acceptance of benefits here, for reasons explained at the end of this subsection).

We might, incidentally, think that a similar worry applies to the dinner club example I sketched earlier in the chapter: imagine that, instead of being a member of the club, I am the partner or friend of a member, and consistently get delicious leftovers. Clearly I have benefited, but this can't mean that I'm under any obligation to cook or give anyone money, and when we consider that (unlike the dinner club) political communities are very often not the kind of thing one can simply opt out of, the worry that the principle as given obligates too many people only intensifies.

For precisely this reason - and indeed, in the context of considering precisely this kind of example - Klosko distinguishes between excludable and non-excludable goods. As we might guess, it is possible or straightforward to provide the former to some and not others, and either impossible or deeply inconvenient (or difficult, expensive, and so on) to provide the latter to some but not others. The goods of the diner's club are fairly clearly excludable goods: if I am not a member of the club then, absent the kind of partner/friend links mentioned in the previous paragraph, I will not receive the goods. Co-operation in the pursuit of excludable goods gives us a pretty simple principle of fairness; if I join a group with the aim of pursuing some excludable good, then I ought to share the burdens or be excluded from the good. As Klosko puts it, "in cases of this sort individuals incur obligations only when they actively seek to attain benefits provided by co-operative schemes" (*ibid.*). This characterisation explains both why I ought to do my share if I am actively seeking benefits, and why being a

lucky bystander - getting the surplus apples, or leftover meals - does not make me obligated in such a way.

Public goods, Klosko argues persuasively, are non-excludable goods: we cannot provide, say, reasonably clean air to citizens A-Y, but withhold it from Z (who lives in the same community). Equally importantly, public goods require that a great many people co-operate effectively in order to exist in accessible form at all. Clean air, extensive public transport, and free healthcare all require a fairly high degree of co-ordination and co-operation in order to be made available, and it is also noticeable that trying to exclude people from the last two results in their immediately becoming more difficult to access for *anyone*. Imagine that there was a surcharge for non-citizen users of healthcare; we would then have to institute systems of identification, rules about residency, and so on and so forth. The principle of fairness, then, requires us to participate in (for example) co-operative labour to administer and supply public healthcare because, i) Such a system is dependent on universal or near-universal co-operation to function, and 2) Excluding non-compliers from healthcare provision would be prohibitively costly (broadly understood).

This will still not do for the purposes of generating political obligation, for the resister still lurks off-stage. Why, they will demand, are we assuming that public healthcare *is* of benefit to them? After all, the interlocutor continues, they know a very fine doctor who will act as their GP at negligible cost, less than would be required of them by the public provision, and so they are neither motivated nor inclined to comply with our demand.

Klosko's response is to invoke presumptive (or presumptively beneficial) goods: goods "necessary for a minimally acceptable life" (1987, p355). Access to public healthcare is a very plausible instance of a presumptive good (whereas we can easily imagine a minimally acceptable life without private healthcare). Agents may thus be compelled by the principle of fairness to contribute to the upkeep of such a good regardless of whether they desire it.

This may seem a little counter-intuitive at first, since it seems that fairness requires contributions from someone who may receive no benefits. An analogy may make the thought clearer and more attractive. Imagine that I have some

rare genetic kink which means that I can tolerate much lower temperatures than any other member of my community, to the point where conditions that would be fatal to my cohabitants merely cause me discomfort. A scheme which provided survival gear for everyone would thus be of no benefit to me personally, but according to Klosko's characterisation of presumptive goods, the principle of fairness still requires that I contribute to the scheme: a minimally acceptable life surely requires that we don't freeze to death (or regularly suffer frostbite, etc.), and goods which enable us to avoid freezing must therefore be presumptively beneficial even if they are as it happens not beneficial to some particular agent.²⁰

There is one other complication to be addressed before assessing whether Klosko's account can do the job we want of a conception of political authority: the distinction between acceptance and 'mere receipt' of benefits. As I suggested above, there are well-known arguments to the effect that the latter cannot generate political obligation; and that the former amounts to an obligation from consent rather than fairness. I argue that given Klosko's notion of presumptive goods, we can (as he, in fact, does) sidestep questions of acceptance versus benefit. Although this move may produce a problem for Klosko, it need not trouble my account (for reasons which become apparent in section 3.4).

This argument can be traced to a famous example from Nozick (1974). As Nozick gives it (p93), the citizens of your community use a PA system for telling jokes and stories, singing songs, and so forth, rather like a particularly local radio station. There is a rota which assigns each person a day when they are in charge of the PA; when it comes to your assigned day, the question is whether your having enjoyed the jokes and songs (for you have) means that you are now obliged to take over and begin spinning tales for the community. Nozick, not implausibly, argues that you cannot be; your mere receipt of the benefits is not enough to generate an obligation for you to discharge. If you have agreed to

²⁰ The analogy also seems to work well as a parallel for the actual motivation behind progressive and redistributive taxation - a giant mansion is not part of a minimally acceptable life, but some secure home is, and so possession of the former does not exclude us from fairness-based obligations to pay some (arbitrarily restricted) tax which funds public-sector housing.

take part in the scheme, then you plausibly are obliged - but only because you have consented, and not because of fairness.

As Simmons (1979) and Klosko (1987) both argued, however, much of the force of Nozick's argument rests on the nature of the goods provided. Simmons provides a different example, and one which seems to foreshadow Klosko's account of presumptive goods. Instead of a PA system, Simmons suggests, imagine a well which must be dug because the community's water supply has become irrevocably tainted (op. cit., pp325-326). Jones votes against the public provision and maintenance of this well, and refuses to take part in its upkeep - he does, however, surreptitiously draw water from it. Simmons' point is that he has thereby accepted benefits and generated, *contra* Nozick, an obligation from fairness and not consent, but we can make a different observation. Clean water is a plausible public good, and Jones is thus obliged on grounds of fair play, not consent, regardless of whether or not he actually *does* draw water from the well (perhaps he could get clean bottled water shipped in, or something of that kind). Simmons later argues against Klosko's principle being fairness-based because of this, but that can be left for the next subsection: for my purposes, the point is that where presumptive goods are concerned, we need not worry overmuch about whether we accept or merely receive those goods.

3.4 The Principle of Fairness and the Three Desiderata

Although Klosko's argument for a principle of fairness as given in the previous subsection is strictly concerned with political obligation of a certain sort - specifically, obligations to contribute to institutions - it is not difficult to widen the scope of the obligation.²¹ If it is unfair to refrain from an action which would contribute to public goods then absent some counter-argument it must also be unfair to *not* refrain from an action which would be damaging to public goods. While political obligation resulting from principles of fairness is most easily understood as implying some claim like "it would be unfair not to contribute to this scheme, and so there will be some rule enforcing contribution on members", there is no reason not to understand it as also entailing the simpler and broader "it would be unfair not to comply with rules which produce presumptive public

²¹ As Klosko notes (1987, p356).

goods". That is, if principles of fairness require rules enforcement, then principles of fairness must be able to give reasons to obey those rules in the first place - and being able to generate reasons for obedience is exactly what we want political institutions to do. With that in mind, we return to the three questions.

1. Does it explain how institutional directives are (or may be) morally binding?
2. Does it explain how these directives get their bindingness from the institution itself (rather than being independently binding directives that the institution simply parrots)?
3. Is it consistent with the relational conception of autonomy that I have defended earlier in the thesis?

My contention is that an account based on principles of fairness is well-placed to give an affirmative answer to questions 1 and 3, but that the bindingness of fairness obligations need not be reliably founded in the institutions; to use a standard example, fairness probably requires that I obey the dictator who issues a directive for citizens to contribute to emergency relief, but this doesn't make the dictator (or, to keep our subjects consistent, dictatorial institution) authoritative. This is not to claim that *no* institution can create fairness obligations: imagine a variation on the traffic rules example where we're all obliged to contribute to some system of public transport that wouldn't exist absent the institution. As we saw in the examination of the NJT, the obligation doesn't trivially 'belong' to the institution just because its particular form is dependent on the form of the directive, but because the institution gives us reasons for action which didn't exist before. My criticism of fairness obligations as following from institutions is rather that such obligations can be generated by institutions which would clearly fail to be consistent with autonomy, and those directives cannot therefore be thought of as binding us *in virtue of the institution*. This being so, a straightforward account of authority as the right to demand fulfilment of institutional fairness obligations won't cut it. I go into this in more detail shortly, but it would be as well to examine the responses to each question in order.

As with the NJT, it should be fairly easy to see how fairness explains the bindingness of an institutional directive. If an institution tells me to do something, and performing that action would be required by considerations of fairness, then I am morally obliged to do so; more modestly and pragmatically, if following an institutional directive seems likely to be the fairest way to distribute (or create, etc.) some good, then I have very strong *pro tanto* reasons to do as the institution tells me.

On to the second question, that of whether the obligations' bindingness is explicable by reference to the institution. At first pass, it seems that fairness-based accounts should be able to make use of something similar to the NJT here - the institution gives us a new reason to perform some action, such as driving on the left or right of the road, which we just don't have absent the directive and/or the institution. The problem is something like this: when a putative NJT-claiming institution gives us a false directive (i.e. it *doesn't* better reflect our reasons) then, contra Raz, we don't have any reason to obey. This gives a pretty tight link between the institution and the bindingness of its obligation; if the institution isn't of form X, the obligation of form (or for purpose) Y, and so on, then we don't have any reasons to obey the directive. A putative fairness-claiming institution, though, can give us reasons - think traffic examples - to obey which didn't exist before, without those being reasons to obey the institution. That is, we might well have good fairness-based reasons to obey Nazi speed limits, but these reasons have got nothing to do with the structure of the institution: it is not set up to serve justice, and any directives it issues which *do* serve justice do so merely coincidentally. Analogously, a non-autonomy-minded institution which issues a directive that does (as it happens) serve the goal of equal maximal autonomy cannot be giving a directive which gets any of its binding force from the nature of the institution. As with the Nazi speed limits, although there might be autonomy-minded reasons to obey the directive, these reasons are entirely separate from the way the institution is set up; the directive cannot bind us in virtue of that institution, and the second desideratum is unmet.

Is this a chimerical worry? After all, the Nazi state would clearly lack authority by dint of failing to respect autonomy (amidst all the other reasons that it would lack authority), so it's not immediately clear that the possibility of institutions

giving directives which we have some reason to obey and some reason to disobey should rule out fairness. We might thus think that the third desideratum offers a way out here: as stipulated, a Nazi state is inconsistent with maximal equal autonomy, so the directive cannot be binding in any case. But the structure of this should, I think, worry us. Take a micro example: some powerful figure comes into my office and lays down a set of rules which will, if followed, result in a fairer distribution of whatever goods and burdens there are to distribute in an office. Let's stipulate that they have no authority in the sense we're concerned with - there is no institutional or procedural reason for me to obey them - but they have the power to coerce us into action, such that their directive will be generally followed and thus be effective. We have fairness-based reasons, one of which is directly concerned with the nature of the directive, to obey this tea-room tyrant, but it seems very odd to say that they give us binding obligations in virtue of their institutional role (or the nature of our institution). As Simmons observes of Klosko's principle, such a claim "would ground political obligation in quite non-democratic contexts" (1987, p270). It is not enough that some particular directives are disbarred from binding on grounds of their inconsistency with autonomy if the issuing institution could still generate political obligation in some contexts despite being anti-democratic (and hostile to autonomy). We wouldn't think that our having institutional and fairness-based reasons to obey Nazi speed limits gave that state *any* authority, and so simply adopting Klosko's principle of fairness is not on the cards for our conception of authority.

It is in response to the third question that fairness accounts give us the piece of the puzzle which is absent in the NJT. While the Razian NJT and its Pre-emption Thesis demand what is effectively domination (as we stand vulnerable to arbitrary interference), a principle of fairness cannot require that we stand in relations hostile to fairness. It is not a simplistic sharing of any and all burdens, after all, but a principle of contributing towards presumptive goods. That is to say, fairness cannot give us political obligations to (for example) participate in unjust schemes, even if our participation would lessen the deleterious effects on others.²² As such, no fairness-motivated account of political obligation should, as

²² Although there might, just about plausibly, be moral obligations of this sort, and there will

the NJT does, command us to obey an unjust or mistaken directive. Something like Klosko's argument from the principle of fairness can avoid making domination a function of obligation, and on this score is therefore consistent with the relational account I endorse.

But the critic will likely still feel uneasy about this. We have stipulated that fairness cannot require domination, so our obligation to obey is no longer wholly divorced from the legitimacy and/or justice of the directive; but it seems now as if the obligation may be divorced from what we desire or judge to be right, and from what might be of benefit to us. Surely this should concern an autonomy-motivated account of authority?

There is a parallel between this question, and the challenge to the relational conception from procedurally-independent desire-satisfaction theorists discussed in Chapter 4. The response here follows a roughly similar line. Recall that when the PIDS advocate argues that preventing someone from satisfying a procedurally-independent desire to be enslaved is an infringement of their autonomy, the relational theorist counters that this is simply to misunderstand what autonomy requires. Autonomy on the relational conception demands not maximal desire satisfaction, but standing in equal maximally autonomy-constitutive relations (where autonomy is understood as having power and authority over one's life). The same applies to the other aspect of slave-holding relations: even if the enslaved agent authentically desires their enslavement, and having slaves is part of another agent's deeply-held desires, preventing the latter from enslaving the former does not infringe either agent's autonomy, because enslaving relations are simply inconsistent with maximal equal autonomy. In other words, "how can I be autonomous if you won't let me do what I want?" is, at least potentially, a question which relies on a mistaken conflation of autarchy and autonomy. Similarly, "how can it be fair to make me contribute to something that doesn't benefit me (or that I disagree with)?" is just to misunderstand what fairness requires.

More excitingly still, what is required for fairness on my account can be given by the purpose of authority (that is, to promote and defend autonomy) - the good,

or at least one good, to be distributed is that of autonomy-constitutive relations, and as previously discussed such relations must include an equality condition.²³ We are committed to something like a principle of fairness insofar as we are committed to autonomy, and we are committed to autonomy insofar as it is a presumptive good.²⁴

Before concluding this section, it is worth considering an objection from Simmons against Klosko's fairness account, and showing that it will not undermine the forthcoming fairness-influenced account of political authority. In 'The Anarchist Position', Simmons argues that Klosko's move from fairness to indispensability does not work; he claims that duties to take part in the provision of presumptive goods "have nothing to do with *fairness* [emphasis Simmons]"; they rest on the moral importance of the goods provided (not on considerations of fair distribution)" (1987, p272). When we have strong reasons to obey some imperative that would provide - say - satisfactory healthcare to others, these reasons are given by the mere fact that healthcare is a (presumptive) good. There would, presumably, be fairness-based reasons to distribute healthcare as equally as possible, but Simmons seems correct in saying that what grounds any obligation to obey here is indeed the presumptive nature of the good and not - or at least not primarily - anything to do with fairness. In other words, when I am coerced into taking part in a system of public healthcare, the reason that I have to obey (and the reason that justifies my coercion) is that healthcare is a presumptive good, rather than that it is only fair for me to take part.

Recall Simmons' example of Jones and the well. I suggested in the last subsection that a Klosko-type argument could avoid questions of whether Jones

23 Recall the discussion in Chapter 4 of the instability of unequal power relations between agents: such relations cannot be justified by recognition respect, and valuing one's own self-direction must ultimately rest on demanding recognition respect for oneself. Since we cannot demand recognition respect for ourselves while denying it to others, unequal relations fail to meet the Strong Authority Condition and so cannot constitute part of a maximally autonomous life.

24 In fact, we're probably over-committed (although not in any pejorative sense) here: committed to equal maximal autonomy insofar as we value our *own* self-direction, committed to maximal equal autonomy insofar as we value agency, and committed to maximal equal autonomy because it is a presumptive good. Notice that the first two apply even if we were to take a sort of sufficientarian line on autonomy as a presumptive good and deny that maximal and/or equal autonomy was required by fairness (although I think that such a line would in any case be mistaken for the kinds of reasons suggested here and in previous chapters), so we have a multitude of good reasons to think that a principle of fairness or some close relative is implied by taking relational autonomy seriously.

had accepted benefits or not (or whether acceptance was consent or not, and so on) by arguing that clean water was a presumptive good, and that Jones was thus obliged to contribute to the well's maintenance regardless of whether or not he actually drew water from it. If Simmons' argument here is correct then, albeit that Jones' obligation does not rest on acceptance or receipt, it doesn't rest on fairness either: Jones is obliged to contribute because we have a moral obligation to help provide others with clean water, not (or not primarily) because his failure to take part in the scheme would be a breach of fair play.

This may be a problem for Klosko. Note, however, that the picture of authority which has begun to take shape over the last few pages need *not* be vulnerable to such an objection, because although there is (to restate the point) something very like a built-in fairness condition with the maximal equal autonomy requirement, fairness is not the initial reason-giving consideration. Rather, that reason is our commitment to autonomy, and as autonomy is a presumptive good, that is a reason we may legitimately expect to apply to everyone. Thus, instead of saying that considerations of fairness demand that you take part in provision of these presumptive goods, we say that the character of these presumptive goods themselves give you reason to take part in systems of provision; and that the equality condition of one of these goods, namely autonomy, demands a fair distribution of presumptive goods. No further justificatory resources are therefore required.

Conclusion

This chapter has argued that the motivation for seeking a workable notion of political authority is, ultimately, autonomy. Without some kind of action-guiding and authoritative institution, we fall squarely into the voluntarist problem and end up with a collection of rugged individualists who may be autarchic, but are not autonomous. Given this motivation, I claim that we should adopt a functional analysis of authority as a promoter and defender of maximal equal autonomy. What it is for an institution to issue an authoritative command is for it to give some imperative which requires collective action (or refraining from

action) and would increase or secure our autonomy-constituting relations.

Three desiderata have been identified: the putative authority must be able to generate (morally) binding obligations, which do not stand independently of the institution, and the justification of and process of enforcing obligations must be consistent with my theory of autonomy. I have argued further that two promising conceptions of political authority, Raz's service conception according to the NJT, and Klosko's account of obligations from fairness, do not meet the conceptual requirements given in this chapter (although each account suggests plausible features which our eventual conception can adopt).

The task for the next chapter is therefore to construct a conception that satisfies the success conditions they imply - a task that cannot, I will argue, be completed by any statist authority.

6.

Authority (II)

Introduction

The previous chapter gave three desiderata for a conception of authority: that it be able to produce morally binding rules, as a result of the authoritative institution's nature, which are consistent with maximal equal autonomy. I have argued that neither Raz's NJT-motivated 'service' conception of authority, nor Klosko's argument for authority from fairness, will quite do the job.

In this chapter, I put forward my preferred conception of authority. After explaining it and showing how it meets the desiderata given in the previous chapter, I pursue two further tasks: defending it against the problems for democratic authority given in the first chapter, and providing a sketch of how we might actually go about meeting the success conditions that the conception provides.

1. The Autonomy-Justified Conception of Authority

In this section, I outline and defend the autonomy-justified conception of authority, whereby institutions can claim authority when they meet two necessary and jointly sufficient conditions:

- 1) The Dependence Condition - Authorities act or effect action in ways which reflect our own reasons
- 2) The Equality Condition - Authorities act or effect action in ways which are consistent with maximal equal autonomy.

In essence, this account borrows the normal justification and dependence theses from Raz and uses fairness (most particularly Klosko's conception of fairness regarding presumptive goods) as a means of explaining why an institution may have authority over those who either deny that such reasons apply to them or that the reasons are reflected by the institution, without this explanation needing to characterise such agents as irrational or acting in bad faith.

It should be made clear that while much conceptual machinery has been borrowed from the various arguments from fairness, it is not *strictly* a fairness principle which does the work here; or at least, not fairness as an independent political principle. Instead, it is the reflexive equality condition of autonomy described in previous chapters. As I have argued (see footnote 4, p171), if we value autonomy then we're conceptually committed to valuing the autonomy of others. This is not a Kantian argument insofar as it need not claim that there's anything *incoherent* about valuing one's own self-direction but nobody else's (although I would tend to think that there is, the argument need not be rehearsed here). Rather, it is a sort of practical rationality requirement. If we are prepared to tolerate or promote only the self-direction of morally arbitrarily-defined subsets of agents, then we stand in a set of relations which have no principled reasons to prevent *us* being given the dirty end of the stick to hold, and so even self-interest provides a good reason to endorse maximal equal autonomy. This being so, I shall refer to the equality rather than fairness principle from now on, in order to signal that the condition is generated internally by considerations of (maximal equal) autonomy, rather than being a separate commitment as it is in the literature on fairness.

1.1. The Components of Autonomy-Justified Authority and the Desiderata of Authority

It may be useful to restate the purpose authority is supposed to serve before laying out how I take the autonomy-justified conception to best meet that purpose. I have claimed that in order to play the role of promoting and defending autonomy, there are three criteria or desiderata for the putative authority to meet: it must be able to generate (morally) binding obligations, which do not stand independently of the institution, and the justification and enforcement of obligations must be consistent with my theory of autonomy. Without the ability to create obligations, the authority is either empty or non-existent; an authority which merely re-asserts free-standing moral obligations is of no use in resolving the kind of blameless disputes which occur within societies; and it should be fairly obvious that failure to respect autonomy is a disqualifying offence for any institution which is supposed to promote autonomy.

Autonomy-justified authority, as I will argue, has the conceptual features required to meet all three of these desiderata. By borrowing the Dependence Thesis' requirement that authorities act or aim to act on reasons which apply to every agent, it can create moral (or morally binding) obligations to obey: if I am committed to autonomy, then I have at least a *prima facie* obligation to cooperate with courses of action which will tend to increase or secure autonomy. Given that autonomy is a presumptive good, we need not be thrown by claims of apathy or active hostility towards it: just as somebody's not valuing physical health may be authentic, rational and perhaps even good *for them* without this making it unjustified to coerce them into contributing towards a public health service, somebody's (authentically, etc.) disvaluing of autonomy does not permit them to act in ways which are hostile to maximal equal autonomy.¹ Importantly, it can also pronounce authoritatively on matters that autonomy is silent on, presuming that such matters exist.² We can easily allow - probably, in fact, the majority of theorists will want to assert - that there are other presumptive goods than autonomy, and so long as the institution correctly reflects our reasons for action regarding these goods then it can still produce binding obligations about them. There will, of course, remain potential problems insofar as it seems possible on a pluralist view (or at least one which held that some values were or could be exclusive) that a policy may pass the dependency test but fail the third requirement, for compatibility with autonomy; but this is a general problem for pluralism, and it would be tedious to go into detail here.

In regards to the second desideratum, the autonomy-justified conception is again successful: only through the creating and sustaining of certain kinds (or broad groups) of institutions can certain obligations be given binding force. This may be in the purely conventional sense, as with injunctions to drive on the left or right of the road, or it may be in the sense that some obligations are only appropriately created or enforced by particular bodies. In general, anyone may coerce anyone into not murdering, but it is inappropriate for the Scottish Football Association to set the rate of contributions to the NHS. We might also

1 See the discussion of presumptive goods and fairness in the preceding chapter for detailed discussion of this claim.

2 As has been alluded to previously, this is something I remain agnostic on within the thesis, but it would be remiss not to at least acknowledge my sole focus on autonomy and sketch how the theory might fit within a pluralist schema.

suggest a third sense of generating obligations in virtue of institutional structure, where neither the convention nor the obligation to find and obey *some* convention applies prior to the authority's policy. Imagine some co-operative public project which is only possible through the operation of a co-ordinating body. This is distinct from cases of pure convention: in those cases, there are strong moral reasons to obey some relevant convention even before the convention is created (that is, we all have independent reasons, based on prevention of harm, to drive on one or other sides of the road, and the convention specifies which side that should be), whereas here both the convention regarding *x* and the need for some convention regarding *x* arrive at the same time.³ Here again, the machinery of the NJT comes in handy. If the authority fails to (fairly) accurately reflect the reasons of its citizens, then it either fails to create an obligation at all - "on Tuesdays everyone must put jam in their socks", presuming that jammy socks are not a valuable component of autonomy - or it creates an obligation that is at best independent of the institution itself - "don't murder" as an injunction of the Nazi state. The equality principle can be seen lurking below the surface here: if a directive laid out by some putative authority fails to cohere with maximal equal autonomy, then it can't be dependent on the institution, because the institution's purpose is maximal equal autonomy.

Finally, the third criterion of consistency with autonomy is guaranteed by the principles of dependence and equality. A policy which fails to promote or defend autonomy is inconsistent with the reasons held by participants within its scope, and the reasons we have for valuing autonomy require that we value maximal equal autonomy.

It will not have escaped anybody's notice that thus far my argument for the autonomy-justified conception's plausibility has consisted largely of assertions. Even if this all looks good, and we are prepared to buy my account of what an

3 Although it's worth noting for conscientiousness' sake, I don't think this potential third type gives us any interesting new ways for the institution to make the obligation *binding*: it's simply that the authority has generated an obligation which is dependent on the authority for its existence as well as its force. In any case, I might just be mistaken here - perhaps such cases are merely non-obvious examples of conventions, in which case the imperative does straightforwardly get its force from being the relevant particular iteration of a general moral requirement.

authority should do and how the autonomy-justified conception fits that account, there is still the question of what substantive features this implies. There is unlikely to be a detailed set of unique practical implications for each requirement of autonomy-justified authority. It is predictable and perfectly reasonable, for example, that similar procedural conditions must be met in order for our putative authority both to act or effect action in ways which reflect our own reasons, and to act or effect action in ways which are consistent with maximal equal autonomy. It would seem odd if a putative authority could require, say, direct democracy in virtue of the first but not in virtue of the second. However, for the sake, or at least hope, of clarity, I have split the upcoming discussion into subsections according to the two success conditions for autonomy-justified authority. This practical discussion will also bear on the conception's theoretical justifications insofar as the ways in which autonomy-justified authority may successfully be claimed will help to illustrate the different conceptual requirements of the Dependence and Equality Conditions.

1.2. The Dependence Condition and Its Substantive Implications

We have stipulated, on something very like Raz's Dependence Thesis, that an institution may only claim authority if it acts in a way which reflects our own reasons, where these reasons include intersubjective reasons to pursue public or presumptive goods. What I call the Dependence Condition provides a necessary rather than sufficient condition on being authoritative in the relevant way. After all, I can demand that someone acts (or refrains from acting) in virtue of autonomy, which they are by hypothesis committed to taking as a reason, without this making me a political authority. It is worth noting that the requirement for authorities to reflect the reasons of their participants is not limited to reasons that they conceptually 'must' have, or what I've followed Klosko in calling presumptive goods. There's no independent motivation (let's assume) for any group to prefer cheese sandwiches over ginger nuts, but if the group *does* prefer cheese sandwiches, then the institution acts illegitimately if it prioritises the production of ginger nuts. For an authority to be able to satisfy the necessary condition of the Dependence Condition, it must reflect our reasons

regarding both presumptive goods and non-presumptive goods.⁴ The immediate task is then to explain how an authority can go about acting in such a way.

In fact, there are two questions buried here. First, how can we create or shape an institution such that it is likely to do this; second, how can the institution actually carry out its task?

It will be helpful here to recall what we have already said about autonomy and democracy in Chapter 1; in particular, the constitutive, respect and equality justifications which may be used to motivate an autonomy-minded theory of democracy. For an authority to act consistently with valuing autonomy, then it must meet these justifications. If it does not act consistently with valuing autonomy then it fails to reflect our reasons and falls foul of the dependence requirement, thus becoming non-authoritative. In short, an authority's being able to satisfy the three justifications is partly dependent on the procedures which determine its institutional composition and which determine how it makes and enacts policy: direct democracy is a requirement of satisfying dependence.

Being able to take part in collective self-determination is very plausibly a requirement of programmatic autonomy, else we are disempowered from making decisions that concern us and our values. We don't have to be "successful" on every occasion - again, such a stipulation would fit only with a desire-satisfaction account of autonomy - but we must not be structurally unable to influence the decision. Such an inability clearly renders us less powerful over the direction of our lives, and also signifies that we are not seen as authoritative over them. This, as previously argued, means that the absolute furthest towards representative democracy we may go is in the selection of *delegates*, not representatives, who vote according to the previous votes of the relevant group

4 In many cases, as suggested later, our reasons regarding non-presumptive goods might just be "we have reasons not to turn to public institutions for the distribution or production of non-presumptive goods", and so the institution best reflects our reasons by making no policies regarding sandwiches or biscuits. However, I don't think this applies in all cases - it seems appropriate to at least consider whether, for example, it would be legitimate to coerce participation in a system which distributes a variety of non-presumptive goods - and there will also be instances of presumptive but unspecified goods where institutions must reflect the actual reasons of their citizens. If we assume that playing football and swimming are both subtypes of a general health and/or leisure presumptive good, and can only provide one of the activities at a particular time, then the question of what form the presumptive good should take should surely be answered by reference to what the citizens in fact think.

(that is: Citizen's Committee A votes to accept or reject various proposals; Delegate A goes off to the national assembly and votes exactly the same way). For a collective decision of any kind to constitute part of our self-direction, we must have a meaningful input in that decision. Selecting one or other charlatan to go and make decisions on our behalf, particularly when we have no method to then influence what decisions they make, does not seem like meaningful input. It's worth noting that, again, a representative democracy with a strict recall system (where the recall is initiated by voters rather than following a top-down dismissal or ejection of a representative) would thus be better placed to claim autonomy-justified authority than a representative democracy of the British 'elective dictatorship' variety. This is consistent with what has already been said about relational autonomy being a scalar rather than binary feature of our lives.

The respect requirement - that we be treated as independent sources of value in virtue of recognition respect of our agency - seems less militantly in favour of direct democracy. It is not immediately obvious that we fail to regard somebody as an independent source of value if they select some other agent to act as their mouthpiece or representative. But look again. When the ability to influence decisions which affect our programmatic autonomy is removed from our hands, *even if we give it up* (authentically, rationally, etc.), then we are so much less powerful over our lives.⁵ It is inconsistent with our being respected as an independent source of value that we not have the powers appropriate to such a source of value; if the Right Honourable Member and myself are equally respected in virtue of our agency, it seems odd that there a structural inequality which gives that member and denies me power to effect sweeping changes to mine and others' lives.

Finally, what about the earlier stipulation that any autonomy justification for a particular system of government must be able to call on equality - that is, that we can only justify (say) democracy through autonomy if democracy produces or affirms the basic equality of agents? Once more, direct democracy is indicated:

⁵ A word on the authority condition here: resisting the temptation to opine, in paternalistic Rousseauian fashion, that people who want to give up authority over their lives aren't to be trusted with that authority, we can instead say that whatever might motivate such a decision is largely irrelevant: whatever the reasons behind my waiving a right to (for example) decide what colour to paint my house, I am still no longer the authority on how my house is to look.

under any system of representation looser than strict delegation, the representative clearly has more power to effect change than do their representees. I do not mean that representatives are more empowered in the faintly trivial sense that I am more empowered to mark essays than my students are, but that such representatives stand in different power relations relevant to their programmatic autonomy: MPs, MSPs and so on are structurally able to effect broad social change in a way that is not true of the average citizen.⁶ Such relational imbalance does not exist in a delegate system, where the delegate's doing things that have not already been approved is directly contrary to their political role (though again, there are comparatively better or worse options in between, such as locally-triggered recall elections within a representative democracy). Note that the equality requirement is, on its own, pretty much silent on whether we should make everyone as powerful as possible, or as powerless as possible, over each other's lives - it's only when we plug in a particular account of autonomy that we can, to call on the broader context of my thesis, decide whether equality demands anarchist communism or right-libertarianism.

The role of the Dependence Condition within autonomy-justified authority is twofold. By asserting that autonomy is a presumptive good, it rules out certain types of organising institution (and background social conditions). And by requiring that authoritative institutions in general reflect the reasons of their citizens, it demands certain types of decision-making procedure as well as certain kinds of background conditions: our reasons are not well-reflected by an oligarchy, even laying aside the presumptive good of autonomy, because an oligarchy is not the kind of institution which is concerned with reflecting the reasons of all its citizens. Neither is a putatively democratic system within a racist and sexist social context; the enforcement of majority racism does not

⁶ Of course, citizens can and do effect change despite this power imbalance - a thought which motivates the anarchist idea of direct rather than political action - but the point stands that it is significantly easier for a member of the political elite to make changes than it is for the average citizen, and that this is unjust because (in part) it suggests differing levels of recognition respect. Taking this to its logical conclusion, we come to the thought that political elites in general are unjust, which seems correct: the trick is therefore to find a form of social organisation which either replaces political with direct action, or makes political action an effectively egalitarian form of engagement. I'm inclined to think that not too much rests on which of these we pick, since the distinction between political and direct action becomes much less clear in a non-statist society.

serve the purpose of producing policies which will tend to reflect the reasons of those citizens.⁷ Even formal equality and a lack of overtly oppressive legislation does not equate to equality of the sort required for democratic procedures to meet the Dependence Condition: if one is a member of an oppressed group in such a society then one *is* structurally less able to influence public decisions, command respect in virtue of one's agency etc., because formal or legislative structures are clearly not the sum or the greater part of social structures.

However, the Dependence Condition does allow for slips or mistakes. If some institution really does, on the whole, produce policies which citizens would pursue their own reasons better by following than by disobedience, then we have a good justification for thinking that any given policy will be of this nature and obeying it. That's true even if there have, in the past, been policies which failed to reflect our reasons. So long as we don't have any grounds to think that a particular policy *will* fail to do so, then its having been produced by direct democratic voting under conditions of fairly strict equality should be a good signifier that it will not fail.

1.3. The Equality Condition and Its Substantive Implications

This condition stipulates that institutions may only claim authority (demand obedience, and so on) when they generally enact policies consistent with maximal equal autonomy. When they *do* act in such a fashion, we are obliged to obey because of conceptual commitments associated with valuing autonomy. Like the Dependence Thesis, the Equality Condition does service at both 'ends' of the authority relation: in order to demand obedience, institutions must meet that condition, but once it has been met we are obliged to obey in virtue of our own commitments. The Equality Condition also provides the sufficiency condition missing from the Dependence Thesis; it is necessary that an institution reflect the reasons of its citizens in order to demand obedience, and it is necessary and jointly sufficient for obligation that a policy would in fact tend to produce maximal equal autonomy.

⁷ In particular: which reflect the reasons, *including presumptive goods*, which those citizens have. As has been argued elsewhere, lots of people endorsing an autonomy-hostile policy does not make it any less autonomy-hostile.

What implications does this condition have for how we might practicably go about putting the autonomy-justified conception into action? Most obviously, it reinforces the previous claims that material equality is a key precondition of institutions being able to claim authority for voting-endorsed policies. In other words, unless and until we live in a radically egalitarian society, something having been voted for does not imply a kind of social consensus which satisfies the constitutive, respect and equality requirements necessary for democratic processes to be justified on the grounds of autonomy.

A nagging thought needs to be dealt with here. We do not, to tender a remarkably obvious claim, live in an egalitarian world. If the only way for institutions to attain any authority is to exist in a revolutionary utopia, then even for a work of political philosophy my thesis is patently irrelevant. Fortunately, the Equality Condition gives us an explanation as to how institutions may claim obedience even within hierarchical and capitalist systems, for it is surely possible for institutions to suggest policies that will tend *towards* maximal equal autonomy even if they start from imperfect positions.

There is an interesting and relevant actual world case here: the recent law changes which expand the scope of marriage as an institution beyond heterosexual couples. We might - as Stonewall memorably if briefly did - deny that trying to co-opt a greater range of people into participating in marriage is of any benefit to those agents' autonomy, because marriage is itself a fundamentally possessive and anachronistic practice, or because the move effectively tried to assimilate LGBTQ agents into a heteronormative narrative, and so on. On the other hand, we might think that there are material benefits to socially-recognised relationship practices which have historically been denied to some oppressed groups, or that challenging and reshaping relationship practices including marriage is required for equality, etc. Depending on what one thinks about epistemic access, the capacity of reformism to stifle radical change, and any number of other concerns, one will naturally be more convinced by one than the other of the positions on equal marriage; but note that one's reasons to obey or not obey will be based on what tends towards maximal equal autonomy in non-ideal circumstances, not on what would be consistent with such autonomy given ideal preconditions.

So: how do we know whether the proposals for equal marriage tend towards maximal equal autonomy? The most obvious answer is the best; we listen to the people whose lives are most affected, and organise around those groups which seem most trustworthy or best serve their role of achieving justice. In other words, questions about what actions we should take, whether as LGBTQ agents, allies, or good citizens, are most likely to be answered authoritatively by democratically organised activist or self-advocacy groups. And here we can ask whether the Equality Condition (in this case) is satisfied by asking whether some directive is consistent with constitutive, equal and respect-grounded autonomy. Could marriage form a constitutive part of somebody's living life according to deeply-held principles? Clearly so. Would marriage being broadened in such a fashion be consistent with equal social relations for others? Again, yes - although in the case under consideration, it is only *consistent* with rather than entailing others' having such relations, and worries about the inegalitarianism of marriage as an institution remain. Finally, is equal marriage consistent with having recognition respect for an agent? The answer should be yes again; regardless of what we think about the value of some practice, there can be no attempt to sanction or forbid the practice on the basis of differential recognition respect - either everyone should, or nobody should, be able to marry.⁸

The three kinds of justification sought for autonomy-motivated reasons suggest that we can very plausibly hold an authority to be valuing autonomy if it endorses equal marriage. The relevant question in terms of the Equality Condition is thus the following: will we be more likely to “act or effect action in ways which are consistent with maximal equal autonomy” if we are for, or against, working towards (or endorsing) equal marriage as state/legal recognition of a slightly broader set of relationships? This is again a substantive question; we cannot, from our armchair, decide which of two putative authorities with competing and exclusive answers is more persuasive. Neither is a more complex answer out of the question, where we believe both that marriage is an autonomy-hostile sham which should be dispensed with and that it is non-trivially worse for folks' autonomy if, in some contexts, they are

⁸ See Chambers and Wedgwood in Brake (ed., 2016) for more on the value or disvalue of marriage and equal marriage.

prevented from pursuing courses of action which would be sub-optimal or inconceivable in ideal conditions.⁹

Given that we are *not* in fact deciding whether to create a relationship practice *ex nihilo* under conditions of social equality, but whether to (incompletely and somewhat arbitrarily) alter existing institutional recognition of relationships under conditions of social inequality, the autonomy-justified conception will say that *if* extending marriage recognition is better than not doing so, we ought to extend it. The equality condition requires that institutions act in a way which preserves or tends towards maximal equal autonomous relations: if we have a case where doing nothing preserves existing inequalities, then it seems plausible to say that even imperfect policies which tend more towards equality are authoritative insofar as they don't shut down or distract from better alternatives.

We would have no reason to listen to anyone who wanted to bring marriage about in our revolutionary utopia, and there are conditions where everyone's autonomy is straightforwardly best served by tossing marriage on the bonfire of other institutions bequeathed by god and government, but in our current actual world it seems likely that it would be both unhelpful and wrong to defy equal marriage legislation, or to work against its introduction. "Don't vote, organise" is a tactical principle, not merely an ideological condemnation, and to deny that *any* reformist change can bring an improvement (even if only relatively and temporarily) is callow and dogmatic.

To put it another way, institutions that cannot claim general political authority may nonetheless be authoritative in certain contexts or within certain scopes - and this seems about right. There is no particular reason to think that obeying a demand from the Anarchist Federation not to eat biscuits in bed will tend towards instantiating or defending maximal equal autonomy, but there plausibly *is* an obligation to abide by a previously agreed-upon plan for what AF members will do during some or other direct action. At heart, our reasons to obey

⁹ Here's another example: the kinds of employment and economic relations presupposed by reformist trades unions are inconsistent with maximal equal autonomy, and often union managers work against the interests of their members anyway. But we still have autonomy-minded reasons which make union directives (about, for example, industrial action) authoritative - although not exclusive of organising action outwith reformist structures - iff we have reasons to think that following such directives will tend to produce more equal autonomy-constitutive relations.

authorities must be derived from reasons we have to secure and further autonomy. If we have reasons to think that some institution's directives are, or suggest, the best option open at the current time, then that institution is authoritative and can claim political obligation from us.

2. Relational Autonomy and the Problems of Democracy

In considering the mountain man problem, I rehearsed at the start of this thesis several familiar criticisms given against autonomy-minded justifications for democracy. They were, very briefly, that:

1. Autonomy motivations for democracy do not survive instances of frustrated self-direction - if I value democracy as a means of instantiating self-direction of a particular kind, the valuing must stop when my self-direction is unsuccessful. Conversely, I will end up endorsing anti-democratic principles to get my own way.
2. Democracy appears to endorse autonomy-hostile and even immoral acts and policies if they garner sufficient votes - if I value democracy because it instantiates or allows for autonomy, for example, I must seemingly endorse a majority decision which was based on authentically-held racist beliefs.
3. The existence of persistent minorities (that is, some unspecified group who for whatever reason always or generally end up being on the losing or unendorsed side of any vote) is inconsistent with autonomy-minded justifications for democracy; if I value autonomy, I cannot approve of any decision-making procedure which systematically deprives folk of their self-direction (or renders their participation in collective decision-making procedures moot).

The total effect of these criticisms, I suggested, was to render autonomy and democracy incompatible - it appeared that we should either ditch autonomy, or democratic authority. Now armed with the relational account of autonomy, and with a sketch of authority as the institutional framework which is required in order to provide maximal equal autonomy, I re-examine the problems and make three claims:

1. The standard formulation of such problems is (implicitly or explicitly) based on a notion of autonomy/self-direction (and thus voting, in this context) as not much more than desire-satisfaction.
2. It is impossible to resolve the problems caused by this desire-satisfaction analysis with solutions which are themselves desire-satisfaction driven. We cannot, for example, avoid problems of democratically-frustrated desires by arguing for a satisfied separate or second-order desire that our desires be achieved in such-and-such a manner.
3. A more plausible explanation for cases where we experience autonomy loss is that our social-relational status has changed and we have lost particular powers or our authoritative status. With a relational analysis in hand, some of the paradoxes of democracy can be addressed.

After laying out the competing analyses, I address each problem in turn, first outlining how the standard (and unsatisfactory) responses to them are rooted in a desire-satisfaction analysis of both voting and autonomy itself and then explaining how the relational view allows us to avoid many of these pitfalls. In the process, I propose the beginnings of a view of autonomy-justified authority's success conditions - that is, how the conception of authority given in the preceding section can be satisfied through democratic procedures.

I argue that under the relational analysis, correctly-constituted democratic procedures will (or at least should) produce outcomes which provide the constitution, respect and equality justifications for autonomy-motivated policies necessary in order to meet the Dependence and Equality Conditions for an institution's being politically authoritative. There may be the conceptual and practical possibility of a directive claiming one or two of the autonomy justifications but failing one of the authority conditions; but any proposal which correctly claims all three justifications *must* be capable of satisfying the two conditions. This is not to say that providing all justifications is sufficient for claiming political authority, as will be shown, but instead that a genuinely authoritative directive will always be consistent with autonomy.

Recap: Voting, Desire Satisfaction, and Relational Autonomy

The desire-satisfaction picture of voting is a relatively straightforward and intuitively plausible one: we vote in order to try and fulfil desires, goals or preferences we hold. In this picture, voting is essentially just a tool to get what we want. It is worth noting that merely having our desires frustrated in this way is not self-evidently an autonomy loss - for voting and losing to affect our autonomy, rather than merely restricting our available choices, something like a desire-satisfaction account of autonomy must also be at work. Indeed, without this underlying presupposition, the idea of voting as desire-satisfaction is relatively inert. What, a critic could enquire, is so damned important about fulfilling our desires?

If we think of autonomy as consisting in achieving our desires (or, more usually, fulfilling desires which have been formed under particular conditions or through particular kinds of reflective procedures), however, the connection between voting and autonomy is clearer.¹⁰ Given material scarcity and reasonable disagreement between agents, some form of democracy offers the best chance of maximising our autonomy. There are a number of ways of cashing this out. It might, for example, be taken to mean that a rational agent would prefer democracy over any other option because it offers the lowest chance of their desires being consistently frustrated, or it could be a return to the maximisation justification.¹¹ However, all of these interpretations look to be operating in tandem with what Elster (1997) calls social choice theory: the idea that voting is something like a purchase mechanism for policies. Importantly, social choice theories suggest not only that people vote *in order* to ‘purchase’ their preferred outcomes, but that the measure of success of a voting system is how well it aggregates desires - that is, decision-making procedures are fit for purpose insofar as they accurately track voters’ preferences, weigh deeper versus shallower preferences, and so forth.¹²

10 For a more detailed account of what I dub procedurally-independent desire-satisfaction autonomy, see Chapter 3, and Colburn (2010).

11 See Harrison (op. cit.) for more on the different ways we might understand democracy as being a maximising or “best-worst” approach to individual autonomy, and the problems with these ways.

12 Elster, op. cit.

On such an account, then, the value of voting for autonomy must be either be linked to its efficiency at satisfying desires, or there must be some feature of the voting method itself which meets procedural preferences we are all assumed to have. The problem cases thus far have all closed off one or both of these options, leaving us with the conclusion that democracy is neither necessarily nor contingently the best option for autonomy.

The relational account produces a different analysis. On this view, although we still vote in order to achieve our goals, the effect voting has is not solely determined by whether or not our desires are fulfilled, and the decision-making procedure which results in the most satisfied preferences (or the weightiest preferences being satisfied, etc.) is not necessarily the best one in terms of autonomy.

Rather, the extent to which our autonomy is protected or promoted (or abnegated) by a vote is determined by the relational changes the vote brings about. To give an example, imagine that a vote is taken on whether or not to run a bypass through my back garden, and I vote in favour of it because I like watching cars go fast. For the desire-satisfaction account, my autonomy is thus increased unless I turn out to have other desires (about noise and air pollution, for example) which are frustrated by the bypass. The relational analysis suggests that how I feel about the matter is orthogonal to its impact on my autonomy. Presume for the sake of argument that this new road will negatively impact my social relations: for example, by resulting in the closure of bus services that allow me to get to and from work easily (thus making it more difficult for me to self-direct). Then, by voting for the bypass I have actually taken part in limiting my own autonomy despite the fact that I have satisfied my desires.

This gives us the beginnings of how the autonomy-justified conception of authority, which aims at maximal equal (relational) autonomy, can be given plausible satisfaction conditions. Where the desire approach is committed to claiming that I am more autonomous insofar as I have more desires fulfilled, and less autonomous insofar as I have more desires frustrated, social-relational theories can argue that while autonomy losses and desire frustration may often

go hand-in-hand, they are not simply equivalent.¹³ With this thought in hand, I now set about the task of re-evaluating the problems of autonomy and democracy.

2.1. Relational Autonomy and Democratic Authority: Frustrated Self-Direction

Consider the earlier case of disagreement over the choice of pub, where we take an informal ballot over whether to go to the Bourgeois Conceit or the Dive Bar, and my preferred option loses out to the massed votes of my friends. On the face of it, this looks like a straightforward (if very minor) autonomy loss, perfectly explicable by the desire-satisfaction account. I had a desire, namely to go to the Dive Bar, and it has been frustrated; assuming that it wasn't a wanton desire - that is, assuming that it coheres with my self-conception and the like, rather than being an inexplicable and unjustifiable urge - then my self-direction has been limited.

Similarly, in the Trident case - where I vote against the renewal of Trident and lose - the desire-satisfaction theorist may feel themselves to be on safe ground. I had a deeply-held desire regarding nuclear weapons, and this desire was frustrated. Surely relational autonomy cannot remove the rabbit from the hat and show that this is somehow irrelevant to my autonomy?

No: the relational analysis gives the same result in the Trident case, but the method by which the result is arrived at offers a far simpler and more intuitive way of rejecting the claim that having to go to a disfavoured pub constitutes a similar failure to instantiate or respect autonomy.

The relations that bound and make up my self-direction in the Trident case are suboptimal in terms of their autonomy constitution: that is, the possibility of nuclear war or accidental irradiation puts my (and, indeed, every other relevant agent's) ability to self-direct under threat. This is true however those relations have come about, whether through dictatorial imposition or majority vote. If the

¹³ Of course, in some cases, the two are intertwined closely enough to make it difficult to separate the impact - if I have a desire to stand in a particular set of (autonomy-friendly, for the sake of argument) relations, and this desire is frustrated, a relatively natural way to think of it is that the frustration of my desire limits my autonomy. Indeed, this is the case, but perhaps misleadingly so: it is only the case thanks to the changes that frustrate my desire being exactly those changes which limit my autonomy. They do not, however, limit my autonomy just *because* they frustrate my desire.

current vote ends up endorsing nuclear disarmament, then those relations are changed for me as well as for everybody else - everybody's autonomy improves, and entailed by that is an improvement in my autonomous status. On the other hand, a decision to renew Trident threatens my autonomy in the same way that it threatens everybody else's, regardless of how they voted for it. In short, the loss of autonomy in the Trident case hinges on the fact that I am powerless to self-direct in a way which is consistent with equal maximal autonomy for others; the relations I stand in are not the relations that allow such direction to occur (or, they are not the relations that would constitute such self-direction).¹⁴ Now, it is likely that the manner in which the decision is taken *does* have some impact on autonomy. Decisions being taken without reference to affected agents, as in the dictatorial imposition of a nuclear arsenal, are not a mark of an autonomy-friendly society. It therefore seems plausible to say that I am slightly better off just for being able to vote on these sorts of things, but the mere fact that I have voted on some decision doesn't mean that the outcome becomes any more autonomy-constituting. If we think that self-direction is *prima facie* important, we should be wary of those things which seem likely to limit our realm of morally permissible self-direction; the atomic sword of Damocles is one such thing, and when we fail to remove it, it is that failure, not the frustration of my desire to remove it, which makes us less autonomous.¹⁵

Drawing on the argument that relations are important because they constitute part of our successful self-direction, and the observation that lots of choices are fairly inert with regards to this, the relational analysis suggests that the pub example changes no important relations for the worse if it even really changes *any* relations.

Firstly, it is not clear what if any relations have been malignly affected by the lost pub vote: I retain my previous power to go to any pub I wish to, and in

14 It is important to note that whether or not the frustration of some deeply-held element of self-conception constitutes a loss or limitation of autonomy is ultimately a question that must be answered substantively rather than formally, as I have argued elsewhere in the thesis.

15 Someone might object that in fact a nuclear deterrent secures our realm of self-direction by protecting it against invasion, etc. It must be said that I find this argument unconvincing in the extreme, but we can make the same point without encountering the objection by reversing the example - imagine instead that the eco-warrior loony left have forced unilateral disarmament against my opposition, and that this will weaken the protection of my realm of self-direction. Again, it's the material change, rather than the frustration of the associated desires, which impacts my autonomy.

addition, I have acquired a new power - I may now go and socialise with my friends. The (temporary) power I have to go to the pub with my friends is constituted in large part by our (unanimous) acceptance of the more general shared wish to go out together. Whether or not I end up standing in a second new relation where I have the power to meet my friends at the pub of my choice, I've still gained this first power of meeting my friends at any pub; and throughout, I retain the only relevant standing I had before the vote was taken. To sum up, the outcome of the vote results in my gaining either one or two new powers, and regardless of what that outcome is, I don't lose any powers at all - so it's difficult to see, if we're engaging in a simple totting-up of relations, how my autonomy can fail to increase; it certainly doesn't decrease.

This explanation, however is vulnerable to the same triviality objection as the desire-satisfaction analysis insofar as it suggests that, assuming I lose the vote, my autonomy is still less-than-satisfactory: I might not be any worse off than I was, but I'm worse off than I could be, and so, as it were, some autonomy has gone missing somewhere along the line.¹⁶ This is an odd conclusion; we are apt to think of autonomy as a matter of being able to pursue deep commitments (although we need not be consciously "deeply committed" to them), and crying foul when we can't go to our favourite pub, buy our favourite fizzy juice or the like just doesn't seem appropriate.

As such, the second option for the relational account is likely to prove stronger. What's going on in the pub case, I argue, just isn't a matter of *programmatic* autonomy - that is, of being powerful and authoritative over the direction of our lives. Programmatic autonomy concerns choosing jobs or careers, where we live, forming and expressing our social and (overtly) political attitudes; the realm of episodic autonomy covers whether to cross the street now or in ten feet, have this or that sandwich for lunch, and so on.

This explains the distinction between being unable to go to my preferred pub and being powerless to prevent nuclear proliferation, or being forbidden from free assembly. In order for us to be able to meaningfully programmatically self-

16 The triviality objection owes its name and its relational application to Gareth Young, who raised the point in response to a draft paper presented at Glasgow University in February 2013.

direct, we must have fairly extensive, stable control over our lives: although there are exceptions (such as, say, a temporary ban on going to a public square because it's prone to bursting into flames), it nevertheless seems generally true that repression of action in the public sphere simply prevents some of the relations required for this, such as being able to stand in defence of or opposition to some programmatic principles without this decision risking one's life or livelihood.

While we may still take this stand under oppressive conditions and thus commit to *authentic* courses of action, part of what it is to be (maximally) autonomous is to relate to others and our community such that authentic courses of action compatible with equal autonomous relations do not carry an unduly high cost. In the pub case, whatever the outcome of the vote there is no change to those relations which concern authentic programmatic self-direction: my episodic autonomy has maybe taken a hit, but the ability to pursue courses of action which further or are consistent with my self-conception remains unaffected, and so does my programmatic autonomy.

2.2 Frustrated Self-Direction, the Three Justifications and Autonomy-Justified Authority

In the pub case above, the proposed directive (“let’s go to the Bourgeois Conceit”) seems to provide all three autonomy justifications. It part-constitutes my programmatic autonomy, it is consistent with my being granted recognition respect as an agent, and it is consistent with all others being able to stand in the same or equivalent social relations.

Crucially, however, it is *not* the kind of decision which can generate institutional obligation, or even be appropriately institutionalised: meeting the three autonomy justifications is necessary, but not sufficient for political authority. Remember that the Dependence Condition requires that a political authority reflects our own reasons for action. In the pub case, the reasons that we have for taking an informal vote are to do with deciding relatively quickly what to do with our evening - reasons which are fairly obviously not best served by putting the question to the vote at the next meeting of the community executive committee. If we were, for whatever reason, to make the decision in that

fashion, then whatever the committee decided would *still* fail to be politically authoritative because the kind of decision being made simply isn't apt for committee oversight: our interests in autonomy are not best served by having (potentially coercive) directives of this sort being put up for general discussion. In the pub example, I am not politically obliged, but it has nothing to do with the frustration of my self-direction and everything to do with the nature of the decision.

If the group discussion was instead a community vote on allowing or prohibiting pubs or pub-like institutions, then the decision would (or would fail to) be authoritative dependent, in part, on whether or not this kind of decision is best made by democratic vote. Plausibly, it is best made in such a fashion; one motivation to think so is that questions about presumptive goods must include the widest possible range of respondents and must be fairly decisively settled. The requirements of the Dependence Condition can tell us when we're obliged to obey institutions, but also when we should consult institutions and how much enforcement power those institutions should have.

It is, in other words, constitutive of our programmatic autonomy that we can make decisions about where to go for drinks, lunch and so forth; but the specific outcomes of those decisions are not (in general) constituents of programmatic autonomy - therefore not presumptive goods - and therefore not capable, even if decided by putative political authorities, of generating political obligation.

To be clear: what happens in a different case, where we vote on the use of common land for some exclusive option - football pitch versus pond - and my preferred option loses out? Here, it seems obvious that the decision *ought* to be authoritative in the sense that it meets the Dependence and Equality Conditions and is an appropriate candidate for institutional decision-making, but we will worry that the institution does not in fact reflect my own reasons. More accurately, though, the decision doesn't reflect *all* of my own reasons. Specifically, it doesn't reflect my reasons for certain non-presumptive goods (let's say football), but it *does* reflect my reasons regarding presumptive goods: it is better for everyone's autonomy, including mine, that decisions about the

use of common land attempt to provide the greatest amount of non-presumptive goods, here determined through voting.¹⁷

The reasons that we have for putting certain questions, most especially those about presumptive goods, through an institutional decision-making procedure are the same reasons that we have for abiding by the decisions which result. At the same time, and to illustrate the constraints placed by the Dependence Condition, the extent to which we can be coerced with regards the football versus pond decision is relatively limited. Perhaps, for example, we cannot be obliged to contribute to the upkeep of the pond (unless we change our mind and make use of it) but can be legitimately coerced into refraining from dumping a harmless but inconvenient solidifying agent into the pond. The idea would be that we (at least by hypothesis) have strong reasons not to interfere in the provision of non-presumptive goods, but no similarly strong reasons to take part in their provision. The work here is all being done by reasons which we've already brought into play with the Dependence Condition: we don't have to be committed to some non-presumptive good in order to have reasons to obey a suitable directive about it.

Given this, the Dependence Condition means that the autonomy-justified authority does indeed generate morally binding obligations as a result of its institutional nature (unlike in the group pub decision case). Presuming that the proposal does not involve driving residents off the land, or threatening folk into voting one way or the other - in other words, presuming that its implementation is otherwise consistent with autonomy - then we have a legitimately authoritative political obligation.

2.3 Relational Autonomy and Democratic Authority: Malevolent Self-Direction

To recap, malevolent self-direction poses two significant problems for autonomy-based justifications of authority. First, malevolent self-direction appears to be inconsistent with the conceptual claim that autonomy is

¹⁷ The mixed justification problem need not be a problem here: we're asking the question "how should we distribute these exclusive non-presumptive goods?", not specifically "what would you personally prefer?", and in this case we can be relaxed about whether the answers reflect self-interest or individuals' principles regarding the distribution of two possible non-presumptive goods.

necessarily respected and its equality condition met by democracy. Second, it also seems at odds with the contingent argument that our interests and self-improvement are best served by democracy.

Viewing the case from a desire-satisfaction standpoint generates an impasse: if we (dubiously) allow that fulfilling or allowing me to fulfil some malignant desire would express respect for my autonomy, the question arises of why my desire for, say, setting fire to cars is taken to be equally or more valuable as others' desires not to be torched on their way to work. Again, the only option that seems open to the desire-satisfaction account is to argue that we must all have some particular desire or set of desires about the way our goals are fulfilled, and that the problem cases violate these. And again, it just ain't necessarily so.

A relational analysis, however, is able to more or less sidestep the problem by arguing that the kinds of proposals which would allow morally problematic courses of action would also make *us* vulnerable. If, to re-state the point, we should respect autonomy because of some value all agents possess which motivates respect, then changing the web of relations such that being an agent is *not* enough to command this respect marks a failure to heed normative importance of autonomy.

A collective decision-making procedure cannot therefore legitimate autonomy-hostile policies on the basis of respect for autonomy where autonomy is understood relationally: even if the policy is overtly hostile only to some entirely notional group, it alters social relations in a fashion inconsistent with the stated motivation for democracy.¹⁸ Imagine, for example, that an overwhelming majority of people in a community vote to limit receipt of social benefits to only those who were born in the community or whose parents were born in the community. If we draw the community broadly enough, so that it encompasses Earth, no existent agent is excluded - nevertheless, the kinds of capacities which are a precondition for autonomous status are, conceptually at least, independent of whether or not one's parents were born on Earth. The collective decision, however, has instituted arbitrary restrictions; given that part of what

¹⁸ This option is not open to desire-satisfaction accounts because it is entirely possible to imagine someone who, knowing that the process which fulfils their desires will render them vulnerable, is prepared to take the risk anyway.

it is to respect autonomy is *not* to take irrelevant factors into consideration when making normative choices, neither the fact that the policy is majority-approved, nor that nobody is *immediately* disadvantaged by it alters the conclusion that the new set of social relations are not justifiable by respect for autonomy.¹⁹

When we apply this thinking to examples closer to home, the conclusion remains the same. Consider the ongoing crusade against (largely imaginary) benefit fraudsters, which purports to be at least partly motivated by increasing folks' autonomy through some ill-defined idea of the autonomy-increasing properties of waged labour. By ensuring that only the worthy poor receive benefits, the rationale appears to run, those who might otherwise let their inherent laziness entice them onto the life of ease provided by £52 a week are motivated instead to find self-direction through employment. But if we consider the impact of these policies, which is to cast anyone who claims benefits as something between a pity case and a con artist, not to mention making it materially more difficult for hundreds of thousands of people to feed themselves, the claims are exposed as hollow (not, admittedly, a particularly difficult trick).²⁰ Even in a possible world where an overwhelming majority of the voting public had directly voted for this proposal, the net relational effect - creating or altering relations which increase economic insecurity, inflict psychological damage, contribute to child malnutrition and so on - is *clearly* not one which respects maximal equal self-direction and diverse self-conception. Thus, although malevolent self-direction in this case may fulfil deeply-held desires about (for example) the kind of society one wants to live in, it cannot be justified by reference to respect for autonomy.

Precisely the same rationale applies with regard to the equality condition. A policy which changes some group or individual's relations on the basis of irrelevant considerations, or considerations which should apply to everyone if anyone, is failing democracy's function as a means of fairly resolving cases where different sets of relations interfere with each other. It cannot therefore

¹⁹ Of course, given that it does not alter any relations between existing citizens and society, the autonomy loss is somewhere between potential and very minor. Nevertheless, it is there.

²⁰ See, for example, Kay, Jordan & Baker (2012), and Beatty & Fothergill (2013).

be justified by reference to autonomy. Return to the single-community-Earth case: if some group of sapient aliens were to unexpectedly arrive on the planet, the policy of withholding benefits from extraterrestrials would constitute an unfair inequality in the democratic process. We cannot coherently justify democracy as a way of paying equal heed to agents' autonomy, and at the same time defend democratically-approved policies which explicitly fail to do so by claiming that their majority mandate means they must be fulfilling the equality condition.

Lastly, the failure of the contingent motivations in this case is explained by the relational approach as a structural matter. If malevolent self-direction is permitted with a democratic mandate, then the defence against domination that democracy is supposed to provide becomes very weak and we have no particular reason to prefer it to a moralistic autocracy or a war of all against all; similarly, it is unlikely that the relations we occupy in the "immoral democracy" case are more likely to bring about valuable traits than relations we could occupy in autocracies or even more hostile situations.²¹

2.4 Malevolent Self-Direction, the Three Justifications and Autonomy-Justified Authority

It should be fairly clear from the preceding paragraphs why proposals which amount to or enable malevolent self-direction will be unable to claim autonomy-justified authority, but it can't hurt to make sure. Imagine, as before, that some policy is put before our committee(s) which will have the effect of forcing unemployed agents to take part in public works projects (on pain of, say, being prevented from accessing some or other non-presumptive good).²² Such a policy will fail to provide any of the three autonomy-minded justifications for democratic authority; and its failure to do so entails a failure to meet either the Dependence or Equality Conditions.

21 The Kantian Gulag, for example, might well prove to be more conducive to valuable moral traits than a society where relations of respect are dependent upon some morally arbitrary feature or features.

22 'Unemployed' might be a slightly misleading term here, since I'm implicitly postulating a society where employment (understood as capitalistic, boss-gives-me-money-for-work relations) doesn't really exist and so to be unemployed is just to have no time-consuming tasks just at the moment, but 'agents of no fixed occupation' is a bit of a mouthful.

Take the Dependence Condition first. Does the punitive proposal reflect our reasons for acting? No. There might be reasons to make *everyone* take part in public works projects (and indeed these reasons would include such proposals being decided upon through democratic voting rather than top-down directives). There might also be reasons for making a particular agent or group of agents perform some task just to keep them busy or to inculcate some attitude or other, such as making a child mow the lawn, or do the dishes, etc.

It is possible, in other words, that actions or programmes of these sorts could be constituents of our autonomy under certain conditions. In the first case, where everyone is equally obliged, the proposals would also give us the equality and respect justifications for their putative authority (presuming that it is consistent with respecting someone as an agent to require their labour or assistance, which it seems to be, dependent on the reasons we give for so obliging them).

But these reasons can't give us any reason to put somebody else's free time in the hands of an institution just on the basis that they don't seem to be doing anything with it. The autonomy-minded reasons that I have for (hypothetically) making my child wash plates are not going to be reasons which suggest the decision should go to a public vote. Equally, the reasons we have for starting and contributing a public project will not stretch to forcing people to take part on morally arbitrary grounds like unemployment. The Dependence Condition, then, is not satisfied by the proposal under consideration: differential coercion of this kind is not constitutive of autonomy, is not consistent with equal autonomy, and does not grant agents recognition respect. It does not, consequently, give us any reason to believe that the institution is acting on reasons which reflect our own.

The Equality Condition is silent on whether or not the proposal should be institutionalised in the first place, but gives an unambiguous answer as to its authority: making some folks do community service but not others is at least *prima facie* inconsistent with relations of maximal equal autonomy. *If* there were some satisfying explanation as to why this *prima facie* inconsistency is actually unobjectionable - if, say, we held that making everyone take part when they had free time was consistent with autonomy, and that all the proposal did was demand greater contributions from agents with more free time - then the

Equality Condition would be met. Again, this is a substantive question, but my suspicion is that such an explanation will not be forthcoming. To give one illustration, it seems likely that making people spend much of their free time working on public projects will severely limit their ability to self-direct in other realms in a way which means the proposal fails to produce equal maximally autonomous relations. Clearly, this is inconsistent with the equality justification for democratic authority.

So the public works proposal given above cannot claim autonomy-justified authority despite its apparent fitness for democratic oversight, because either it's inconsistent with our reasons regarding autonomy (the Dependence Condition), or it's inconsistent with agents standing in relations of *equal* autonomy (the Equality Condition). Malevolent self-direction, whether in the form of setting kittens alight or forcing punitive labour on the unemployed, cannot be justified by autonomy and therefore cannot be justified by democratic majority.

2.5 Relational Autonomy and Democratic Authority: Persistent Minorities

The failure of the desire-satisfaction analysis is not so much in explaining the problem of persistent minorities as in resolving it. In a hypothetical society of a hundred people, where thirty are bound to lose any vote on a particular matter of importance to programmatic self-direction, it is fairly easy to describe the breakdown in terms of frustrated desires: not only is there some deeply-held goal which cannot be achieved now, but it cannot be achieved at all under the obtaining conditions. Further, it seems likely that folk do in fact have a negative desire of the sort "I do not want my desires to be irrelevant to collective decisions" which is being frustrated into the bargain.

Plausible enough so far: democracy cannot instantiate or respect self-direction or our self-conception if fulfilling these conditions is a matter of satisfying desires and we are structurally blocked from doing so, and as suggested above, this structural exclusion also looks to violate the equality condition. But posing the problem is only half the task, and the desire-satisfaction analysis cannot provide an answer.

Just as in the previous cases, the only way for a desire theory to try and deflate the autonomy loss is by pointing to some higher-order desires regarding decision-making procedures, and for the reasons provided previously, this strikes me as unsatisfactory. Indeed, with regards to persistent minorities our desires are if anything more likely to militate *against* a process which puts us in a disproportionately weak position because of prevailing social or political trends.

The relational account is unable to explain away the autonomy loss by casting it in terms of changing relations - whether our hypothetical minority are thought of as having deep desires frustrated, or of being put in relations that do not instantiate programmatic autonomy, they are still patently not self-directed, and as such there is no relational magic wand to be waved at the problem. What can be done, however, is to relocate the problem by recalling why it is that we are supposed to value democracy.

If democracy is to be justified by reference to autonomy then, at the risk of obviousness, it must be an appropriate kind of process and it must range across an appropriate realm - to put it in relational terms, it must (conceptually) instantiate and respect and (contingently) tend to produce maximal equal autonomous relations. That is to say, there are some kinds of proposals - again, arbitrary restriction of freedom of assembly is a useful example - which fail the conceptual and contingent tests regardless of how many votes they garner or how many desires they satisfy. Looking to autonomy to justify democracy cannot work if the sorts of proposals which the process purports to legitimate are inherently hostile to autonomy-constitutive relations. The problem in the case of persistent minorities is not so much that some group of people keep getting their desires frustrated, but that their minority status is enough to make their ability to programmatically self-direct vulnerable.

This, incidentally, is why desire-satisfaction theorists cannot adopt something like the relocation approach and insist that democracy is only justified by autonomy if it respects programmatic self-direction. If the relevant programmatic desire was something like the setting up of a fascist state, or ethnic segregation, then the minority's continuing frustration is no failure to respect or instantiate autonomous relations, but it is, clearly (just as was clear

in the Trident case), a denial of deeply-held desires. That is to say, being an active fascist might well form a subjectively-crucial part of someone's self-direction; without being embedded in a society which reflects and expresses this ideology, they are in a real sense not the person they would like to be. Denying them this opportunity need not count as an abrogation of autonomy on the relational account, for their desires are fundamentally incompatible with the normative basis of autonomy's value and with the kinds of relations which are most evocative of this value. However, the desire-satisfaction theorist must say that so long as the relevant desires have been formed under conditions of procedural independence, the fascist is less autonomous in a democracy than in a fascist state. This looks to me like a very good reason to be sceptical of the desire-satisfaction account, and a correlatively good reason to endorse the relational analysis.

2.6 Persistent Minorities, the Three Justifications and Autonomy-Justified Authority

An institution which fails to provide to agents relations that partly constitute programmatic self-direction, on the basis that those agents are members of a minority group, will obviously fail to have autonomy-justified authority. If, for example, some organising body persistently enforces rules which prevent a group from wearing (or not wearing) certain clothes, or hairstyles, etc., then that body is failing both the Dependence and Equality Conditions. It is important to be clear, however, that it is not merely the fact that the institution has frustrated the agents' self-direction which causes this failure of authority. Instead, there are two steps the institution takes which result in the failure.

In the first instance, this kind of decision is not apt for general oversight; it is not the wearing or not of some item of clothing which constitutes our autonomy, but of standing in relations such that our sartorial choices - whatever their motivations - are our own business without strong countervailing reasons. We can imagine that racist or sexist slogans on T-shirts would be one of the exception cases, dependent perhaps on whether we think that folk have the right to be out in public without effectively having hate speech directed at them,

but that isn't the crucial issue here.²³ As we've given the example, there aren't any of these possible countervailing reasons. It is simply that the institution has decided to raise the question of whether such-and-such an item of clothing is generally permissible. It therefore immediately, and obviously, fails the Dependence Condition: we don't have any general reasons to put up this area of our lives for institutional approval.²⁴ Room to decide for ourselves how we dress or wear our hair is a constituent of autonomy, and for obvious reasons granting someone recognition respect requires that we give them (at least *prima facie*, as mentioned before) this room. Policies like the one under discussion are thus doomed from the off, because how we dress is not something that should in general be determined by institutions - we have no reason to put the matter to a vote, let alone abide by the outcome. So, the Dependence Condition goes unmet here because any institutional directive about hair length fails to reflect our reasons - in fact, such a directive would be in direct opposition to what our reasons suggest about authoritative directives and institutions.

The institution, then, ought not to be issuing directives on how people dress in the first place (excluding some fairly specific cases). Furthermore, and displaying some similarities with the example in Section 2.2, the institution certainly has no business placing agents under restrictions because of their membership of a minority group. Not only is the institution trying to coerce agents illegitimately, the coercion would be unjust even if it met the Dependence Condition; it is clearly false to say that a policy which had the effect of limiting agents' self-direction on the basis of morally arbitrary factors is consistent with, or tends towards, maximal equal autonomy.

23 Notice that even in this case, we'd probably think that the institution cannot rule on what we wear at home - our reasons for particular endorsements or prohibitions differ across different cases, and the negotiation of interactions between agents in common spaces is sensitive to different reasons than the interactions between agents in private (in the natural language sense).

24 In contrast, and continuing the thought in the preceding footnote, we have reasons to negotiate with, say, flatmates about this kind of decision. Negotiations of this sort can't have the same sort of coercive power, however, because there will be context-sensitive reasons that don't apply to general policies. For example, if two of my flatmates object to a T-shirt on grounds of personal taste (that is, they genuinely revile it, but it's not offensive or oppressive), I have good reasons not to wear it around the flat. But they cannot *force* me not to, and they cannot prohibit me from wearing it out and about, because the sorts of reasons that apply in flatmate negotiations won't (or won't always) apply outwith that context, and don't endorse coercion even within that context.

In sum, we have no reason to put up our choice of clothes as a matter for general oversight, so the policy fails the Dependence Condition for institutional authority. And even if we did have such a reason - such as the potential racist T-shirt case - a policy which makes only some (morally) arbitrarily-defined group's clothes choices a matter for general oversight would fail the Equality Condition.

One further problem should be addressed here. As suggested in the previous subsection, my account will say that denying fascists the opportunity to endorse fascism at the ballot box is no limitation of their autonomy.²⁵ How then can an institution claim autonomy-justified authority for a decision which non-trivially limits some group's ability to effect social change, on the specific basis of that group's features?

Take the Dependence Condition first. When an institution refuses to put fascistic proposals to the vote, does it reflect our own reasons? Yes, and despite the apparent counter-intuitiveness of the claim, this includes at least some of the reasons fascists will have. Specifically, we all have reasons from the value of self-direction to endorse a decision-making process which will not render our programmatic self-direction vulnerable to arbitrary interference; since there is no iteration of fascistic thought which will not render us so vulnerable, we thus all have reasons to support a kind of 'no platform redux'.²⁶ Someone may authentically endorse fascist government, but as we have already seen that authentically endorsing some course of action does not equate to that course of action constituting autonomy, this is not an objection for my theory. An agent who denies that such a government would render them vulnerable, meanwhile, is simply incorrect. Power wielded for morally arbitrary reasons admits of no non-arbitrary restrictions: once you have pissed in the kettle, you cannot expect

25 Again, and just to be clear: I am not arguing that we should keep fascist proposals off the agenda but let the fascists organise in the streets. Here, as elsewhere in the thesis, I focus somewhat artificially on voting as *one* method by which we organise our society, but this simplicity-motivated focus should not be mistaken for an implicit claim that only through the ballot can we make legitimate policy and organisational decisions.

26 'No platform' being the anti-fascist principle that fascists ought not to be granted a platform for organisation - whether through spreading their views in public debates (which also serve to legitimise fascism as part of the political norm), holding demonstrations, and so on. It is important to note that no platforming is not the 'censorship' of fascist ideas, but the mobilisation against those ideas being put into action. See <http://flag.blackened.net/revolt/ppapers/fascists.html> (among others) for a fuller explanation of no platform in principle and practice.

your tea to taste the same. The anti-fascist policy under observation, then, is one that anyone with an interest in autonomy - which is all of us - has a good reason to defend that autonomy. Preventing fascists from coming into power is necessary to defend autonomy, and so an institution which gave fascists no platform would, at least in this instance, pass the Dependence Condition.²⁷

The means by which such an institution would meet the Equality Condition are similarly straightforward. Is a policy of hostility towards fascists consistent with maximal equal autonomy? Yes - it is not arbitrary to treat agents and groups differently according to substantive differences in their outlook, dependent on what the effect of those substantive differences are. That someone is a fascist is not arbitrary in the way that their being white, or gay, or elderly, is, or even in the way that their being a meat-eater or a morally flawed individual would be arbitrary here: the things (or at least some of them) that they endorse and value as important are simply incompatible with others being equally as autonomous. It is not therefore a failure of equality to prevent them from bringing such inequality to pass.

Close of Section 2

I have argued in previous chapters that while desire-satisfaction theories can account for some of the problems of democracy, they cannot coherently solve these problems. There are also independent reasons to be suspicious of a pure desire-satisfaction analysis: the fulfilling of malevolent desires in particular

²⁷ It might be thought that (as Robert Cowan has suggested) this reasoning implies that we should no platform not just fascists, but also liberals, socialists, and indeed anyone whose beliefs include a commitment to the state - after all, their values are (by hypothesis) inconsistent with equal maximal autonomy. There is more to this objection than can be satisfactorily dealt with here, but I think that there are two key differences between no platforming fascists and no platforming any statist. First, although statism is indeed hostile to maximal autonomy, the effect of shutting down all statist-friendly political debate rather than engaging with it is also inconsistent with autonomy - the cure is likely to be worse than the disease. Secondly, and as suggested in the last footnote, what makes the toleration of fascism particularly dangerous is that where liberals (for example) organise, they do not *tend* to bring with them physical attacks on oppressed minorities. Fascists, on the other hand, invariably do. It's not 'just' that fascism as an ideology is especially inconsistent with autonomy; it's that letting fascists organise and giving them the veneer of mainstream acceptance (thereby encouraging them further) is directly inconsistent with autonomy. Thus, although the basic theoretical incompatibility of fascism with autonomy provides an argument for why no platformism doesn't fail the Dependence Condition, it is the specific nature of the fascist threat which gives us the "all things considered" justification for no platforming - here, the disease is worse than the cure.

looks capable of being waved through so long as the desires are procedurally independent.

A relational approach, by contrast, can not only explain the apparent autonomy loss in the problem cases just as well if not better than the desire account, but also go some way to resolving the problems. The key matter for justifying democratic authority on a relational account of autonomy is the effect of voting on the social relations that we stand in (or the relations revealed by putting a matter to the vote in the first place), and an authority which meets the Dependence and Equality Conditions will not malignly affect relations which are constitutive of our programmatic autonomy.

3. Some New Problems for Autonomy-Justified Authority

I have argued so far that the relational analysis of autonomy and the autonomy-justified conception of authority resolve the three familiar problems for democracy of frustrated self-direction, malevolent self-direction and persistent minorities. Conceiving of authority as a matter of institutions meeting the Dependence and Equality Conditions, and of autonomy as matter of relational standing rather than desire-satisfaction, allow us to show that the kinds of policies which would malignly affect our relational standing will not be made authoritative by democracy; but, still, that democratic endorsement of some policies *do* give us reasons to treat them as authoritative. In this respect, the autonomy-justified conception has done the job required of it. It can point to autonomy as an explanation for how properly-constituted direct democracy may generate political obligations, and evade the conclusion that respect for autonomy requires the abandonment of political authority.

But it would be premature to put out the jelly and ice-cream yet. Before endorsing autonomy-justified authority, there are two worries specific to it which must be addressed: that it gives a problematically perfectionist view of authority via the Dependence Condition, and that the Equality Condition is redundant and therefore uninformative about authority.

3.1 The Dependence Condition and Perfectionism

As is obvious, the account of autonomy that I've given in the thesis is broadly perfectionist insofar as it claims that there are substantive criteria for an agent's being autonomous - so the voluntary, authentic (etc.) slave is not autonomous, while the agent whose authentic desires for submission to or governance by another are frustrated is autonomous despite themselves. Although perhaps less obviously so, the conception of authority given here is also a perfectionist one. Two proposals can have passed through exactly the same procedure, been approved in exactly the same way, and so on, and yet not be equivalently authoritative because of substantive differences in how they (fail to) reflect our reasons.

This means that my account must deal with the following familiar objection: what is there to defend non-malevolent but apparently valueless self-direction against directives which aim to produce valuable self-direction through coercion? We can put it the other way around, too: if the Dependence Condition requires that authorities act to promote or defend autonomy-constitutive relations, then do we not end up with an account of political obligation which says that I may be coerced out of (e.g.) counting blades of grass and into evening classes?

The particular nature of the problem as it applies to the Dependence Condition should be made clearer. When an institution claims authority on the basis that it acts on reasons which reflect my own (in other words, according to the Dependence Condition), these reasons include presumptive goods like autonomy. With a substantive account of autonomy, therefore, it seems possible that an institution can coerce me into doing something that I do not value or hold as part of my self-conception, on the basis that doing that thing will come to form part of my valued self-direction. But this just looks like exactly the kind of false consciousness, 'forced to be free' approach that we have already rejected. If it turns out that the autonomy-justified conception of authority commits us to this, then the conception is worthless after all and the mountain man is vindicated.

Fortunately, a response is possible, and helpfully brings out the internal constraints which the Dependence Condition places on coercion. The response hinges on two points: the authority's purpose is to promote and defend

autonomy-constitutive relations, *not* to force people to stand in them; and we have no reasons to give authorities the kind of intrusive powers that would be necessary to conduct this kind of objectionable perfectionism.

Take the first part of the response. When an authority makes it so that, for example, it is feasible for agents to become teachers, or lumberjacks, or whatever, it is serving its purpose (at least in this case) because it is expanding the scope of our programmatic self-direction. We would not think that it had failed to be authoritative if, as it happened, nobody's deeply-held values led them towards being a lumberjack - the institution reflects our reasons for attaining the presumptive good of autonomy insofar as it makes being a lumberjack possible, not insofar as it causes lots of people to be lumberjacks. While counting blades of grass is perhaps an implausible candidate to be central to an agent's valued self-direction, insofar as it is possible then we also have reasons to think that its being permissible does form part of the presumptive goods authorities should assist in providing. To emphasise, it is not that we all have reasons to count blades of grass and therefore the authority should reflect these reasons, but that we have reasons for authorities not to interfere in our pursuit of self-conceptions and central values. It is thus a mistake to describe counting blades of grass as 'valueless' self-direction; it might be silly, or implausible, or what have you, but insofar as it forms part of somebody's self-direction and doesn't interfere unjustly with anyone else's, it is indeed valuable.

It would be absurd to expect that the Dependence Condition requires authorities to reflect every single reason for action that every agent has; it rather requires that authorities act in ways which we all have reasons to endorse, and one such way, plausibly (and as argued in Sections 2.1 and 2.3) is to give us room to pursue esoteric or iconoclastic life plans. Now, it is possible that authorities will in fact be justified in prohibiting something harmless on the basis that there are other goods to be pursued: maybe our grass-counter is sitting in the middle of a popular football pitch. But here, of course, the prohibition (or rather, lack of priority) has nothing to do with making the grass-counting agent more autonomous and everything to do with co-ordinating public resources such that any given agent is likely to be more autonomous. It is not for the agent's own good, except in the very broad sense that we all benefit from living in a society

which tends towards maximal equal autonomy - and in any case the authority *still* has a reason to try, if at all possible, to provide the grass-counter with a different patch of land to contemplate.

The second part of the response is negative rather than positive. In addition to having reasons which entail that authorities provide space for unusual lives, we have strong reasons *not* to allow institutional or general oversight which would be likely to make morally permissible self-direction more difficult. Imagine the kind of institutional machinery required to coerce our luckless Stoic into evening classes - some group has to decide that his harmless self-direction is a proper matter for their oversight, then they have to force him to do something which is neither morally obligatory nor valuable to him (nor required for the instantiation and defence of equal maximal autonomy-constitutive relations) and, presumably, they must stand ready to punish him if he resists. This is not, I suggest, a setup which seems consistent with a society which values self-direction on the basis of respect for agency, and if we were to find ourselves in such a society we would all have strong reasons to destroy or radically alter the offending institution. In other words, considerations of autonomy motivate us strongly *against* giving institutions the ability to interfere with our self-direction unless that self-direction is itself inconsistent with maximal equal autonomy. Incidentally, this will also limit the kind of coercion that we can employ if the grass-counting man does need to be removed from the football pitch - in fact, the permissible coercion is probably limited to exactly that removal and no more, although by the same token it's permissible for the impatient players to remove him rather than waiting for some council officer equivalent.

The Dependence Condition does therefore imply a substantive and broadly perfectionist rather than procedural and content-neutral account of authority, but not problematically so: it is met when directives reflect our reasons regarding presumptive goods, and hence will be sensitive to whether some proposal is in fact consistent with maximal equal autonomy, relationally conceived. The Dependence Condition does not legitimate coercion which forces us to pursue "more" autonomous lives, however; in part because this is a misconception of autonomy as being constituted by fulfilling lots of authentic desires rather than relational standing, but also because exactly those reasons

which justify political authority serve to set the scope of that authority. We just don't have any reason to give an institution the authority to forbid us from or punish us for grass-counting, and so no institution attempting to claim such authority can meet the Dependence Condition.

3.2 Is the Equality Condition Redundant?

The other immediate worry for autonomy-justified authority is taxonomical or methodological. The Dependence Condition, as I have stated it, requires that legitimate authorities act on reasons which reflect our own, including presumptive goods. If one of these goods is autonomy, specifically the relational conception of autonomy given earlier in my thesis, then it may appear that the Equality Condition is redundant - authorities are already required to act in ways which reflect our interest in maximal *equal* autonomy, so what extra work is being done by the stipulation that authorities must act in ways consistent with equal autonomy?

The first way to respond would be to accept that the Equality Condition is just one particular gloss of the Dependence Condition, in the same way that the Formulae of Humanity and Universal Law are supposedly two ways of cashing out the same Categorical Imperative. In this case, the objection is not so much that the Equality Condition is wrong but that it's just unnecessary: all that is required to generate political obligation is that an institution meets the Dependence Condition. It might be helpful to draw out the Equality Condition as a particular implication or sub-condition, but ultimately political authority is just a matter of satisfying the Dependence Condition. This would, of course, be inconsistent with what I have already said, namely that the Equality Condition acts as a sufficiency condition in tandem with the Dependence Condition: if the Equality Condition is just a gloss, then it cannot provide different logical success criteria to the Dependence Condition.

But this is not a very satisfactory answer. Apart from anything else, it risks circularity - the Equality Condition must be different because it plays a different role in legitimate authority, and it plays a different role because it is different (and it is different because it plays a different role, and round and round we go).

I must, then, show how it is that the Equality Condition gives us sufficiency conditions which are not provided by the Dependence Condition. My argument is that a directive's meeting the Dependence Condition just shows that it *could* be authoritative; it is only when something meets both conditions that it *is* authoritative. There is a practical bent to the Equality Condition which allows it to play this role, as can be seen in the following example.

Return again to the case of the agent who wants to sit and count blades of grass. As we have seen, there is no in-principle reason why this should fail the Dependence Thesis - it is consistent with relations of equal maximal autonomy that the agent should self-direct in such a fashion. As such, an institutional directive of the form "Park X is primarily for the counting of blades of grass, and that activity should be prioritised there" can plausibly meet the necessary conditions for authority. But this is - as in fact became obvious in the previous discussion - not enough to actually make the directive authoritative. Why? Because it may be that in the actual world, we do not have the available resources to enable everyone to self-direct in the (morally permissible) ways that they want to. Perhaps, as suggested in Section 3.1, a choice must be made between two activities which are both consistent in principle with maximal equal autonomy, and both exclusive (or more precisely, exclusive at place p and time t). The question then becomes one of whether there are differing costs of implementation, and which of the two activities is in practice more consistent with, and more likely to result in, a greater range of equal autonomous relations. That is, the issue which the Equality Condition considers is "which of policy A or policy B is the most likely, given pertinent empirical information, to serve the end of equal maximal autonomy?". This will include considerations of implementation that may not be taken into account by the Dependence Condition because such considerations may apply solely in practice rather than in principle. It is plausible, for example, that a policy of not interfering *at all* with folks' morally permissible self-direction would meet the Dependence Condition (indeed, it is a very attractive policy in principle) but would fail the Equality Condition because, in practice, the result of total non-interference would be the establishment or maintenance of informal power hierarchies which are in principle and in practice inconsistent with equal maximal autonomy.

Conversely, a policy which suggested strong perfectionism without a coercive bent - we try *really hard* to convince the grass-counter that he would be better served doing something else, but we don't manipulate or force him into doing something else - would fail to be authoritative, even if it was not objectionable. This is precisely because any kind of obligation it came up with would be impossible in practice for the institution to enforce: an authoritative institution creating a policy with, by design, no associated coercive powers and hence no purpose to claiming political authority. It would, to rephrase the claim, be attempting to generate an obligation with no obligatory acts; and if it tried to make some acts obligatory, the purpose of the policy would be defeated (because the policy is supposed to be non-coercive).

The Equality Condition is therefore not redundant because it completes our principle of legitimate political authority by telling us something about which the Dependence Condition remains silent: conflict between claims. It, first, allows us to individuate between two (in principle) equally permissible but exclusive proposals or directives. It also serves to account for the illegitimacy of policies that are (again) in principle permissible but would in practice, and for contingent reasons, be inconsistent with the stated purpose of the institution in question. We can think of the Dependence Condition as catching policies which institutions may never endorse and offering directives which could plausibly be authoritative, and the Equality Condition as telling us what institutional directives we must follow, as well as catching directives which some institutions may issue - but not *this* one in *this* place.

Conclusion

I have argued in this chapter that an autonomy-justified conception of political authority is implied by the foregoing thoughts about the need for and purpose of any political authority. Under this conception, institutional directives are authoritative when they meet the Dependence Condition (when an authority may claim to "act or effect action in ways which reflect our own reasons (these reasons including autonomy as a presumptively-desired good)" and Equality Condition (when the institution will in fact "act or effect action in ways which

are consistent with maximal equal autonomy”). Further, I have argued that a direct democracy, under conditions of equality, is capable of meeting both these conditions and thus claiming legitimate authority. Some points should be clarified or re-asserted before moving to the final chapter.

Firstly, some kinds of proposals are simply not appropriate subjects for legitimation by reference to autonomy; you cannot make a fundamentally autonomy-hostile policy into an autonomy-promoting one by covering it in ballot papers. Given that we are arguing from autonomy to democracy - given that democratic authority is justified by reference to how well it instantiates, respects and promotes autonomy - it is counter-productive to try and argue that a proposal which patently fails to instantiate, respect or promote autonomy can be made to do so just by having majority approval. The democratic process, then, must not be taken to magically legitimate autonomy-hostile proposals just by virtue of the fact that it is supposed to serve partly as a method of autonomous self-direction.

Secondly, for voting to have a meaningful impact on our relations, there must be equality at more than the ballot box. If, for example, a community passes a law requiring that medical treatment be free at the point of use, it will not put me in a notably better position if there is no relatively easy and reliable way for me to access that treatment in the first place. The merely formal change, that I may now theoretically go to a doctor without taking my wallet, does not constitute enough of a change in my powers of self-direction to argue that the proposal has done much to increase my autonomy - for that to occur, it must be the case that there are other changes (to public transport, or expanding the healthcare system, for example) which mean that it is now in fact easier for me to seek medical care.²⁸ To justify democracy through autonomy, then, we have to think of democracy fairly broadly as a whole set of social institutions, of which voting is one, but which also includes ideas of freedom of expression, freedom from inequality, and so on that are best served by a direct democratic decision-making process.²⁹

²⁸ The corollary of this is that, had my situation been different (had I lived right next to a hospital, for example) the change *would* have had a fairly dramatic impact on my autonomy.

²⁹ This being so, it invites the question of why I have more or less equated democracy with the

This all militates, as I have been muttering consistently throughout, towards a radically egalitarian and participatory view of political authority. It should be fairly obvious that such a view will be far more at home in the communist anarchist camp than as, say, a contemporary liberal account of authority. What I have not as yet done, however, is to make clear why I take it that the view is not just consistent with, but demands, an anarchistic organisation of society. In the next and final chapter, I recapitulate the argument that states can never grant us recognition respect in virtue of our agency, and offer a sketch of how an anarchistic society might be organised in light of what I have said about autonomy and authority.

narrow conception throughout this chapter. The answer is twofold: for pure simplicity's sake, and also to draw out the point that autonomy-minded justifications for democracy must ultimately be concerned with more than formal or procedural matters.

7.

Summary and Concluding Remarks

1. The State of Play

At the start of this thesis, I presented the problem of the mountain man: the individual who argues that a concern for autonomy is inconsistent with the existence of legitimate external authorities. The mountain man, I suggested, was a problem for autonomy theorists in general, but particularly for autonomy-minded anarchists. Given that anarchists deny the legitimacy of states, and are historically unfriendly to institutional coercion of any sort, the net effect of the mountain man's challenge seemed to be that anarchists must drop one of autonomy or authority - either the value of autonomy is not a good reason to reject state authority, or the value of autonomy is such that only an explicitly contractarian and voluntarist account of authority is plausible. This latter view, I argued, essentially amounts to accepting the mountain man's claims, and given that communist anarchists repudiate the sort of Nozickian minarchism or free-market libertarianism entailed by these claims, autonomy-minded anarchists now have a trilemma.

First, accepting autonomy and authority but not anarchy suggests a familiar account of something like social democracy or left-liberal democracy: there can be legitimate states, and some directive's being a law does provide a strong *prima facie* reason to obey it. It should not be difficult to see why this is unappealing from an anarchist perspective.

The possibility of autonomy and no legitimate political authority must lead to individualist anarchism, which we have already rejected on the basis of its unattractive fetishisation of autarchy and self-sufficiency.

Finally, we have the curious third option of authority and anarchy at the cost of autonomy.¹ Given that one crucial goal of my thesis is to find a way of making

¹ This last claim seems particularly odd, but is not completely inconsistent with the historical goals of anarchism: perhaps, in fact, communist anarchism is entailed by a concern for equality which rules out having autonomy as a primary value (some of the feminist critiques

autonomy and authority consistent with each other, this is less a plausible option for achieving that goal and more a way of explaining why it is impossible to do so.

If, as the trilemma above suggests, the only means of reconciling autonomy and authority entails the legitimacy of states, then autonomy-minded anarchism is incoherent - and, again, autonomy has and continues to be a key value in much anarchist thought.² The task has therefore been to show that autonomy and authority are compatible with each other, and that a legitimate non-voluntarist political authority is not inconsistent with communist anarchism.

My response has involved three elements. First, I made a methodological claim that we do best to delimit the concepts of autonomy and authority functionally, according to the role the concepts play in our political discourse and practices. Second, I gave an argument to the effect that the most plausible way to conceive of autonomy is relationally, so that we understand autonomy as being in large part constituted by the social relations that we stand in. Third, I defended an account of authority as a means of promoting and defending particular values - in this case, authority as an institution which allows us to promote and defend autonomy-constituting relations.

In Chapter 2, I argued for the functional analysis of autonomy and authority, borrowing the approach from Craig and Lane. I suggested that the concept of autonomy best captures the reasons we value and practice self-direction if we understand it to require that we be powerful and authoritative over our lives: the autonomous agent is competent to choose courses of action which reflect her values and principles, and her having chosen them is a *prima facie* reason for her to be empowered to carry these courses of action out. Such a concept evades the charges of unpalatable individualism and misdirected valorisation of

of autonomy as a concept which were examined in Chapter 2 could also lead to a position much like this). I think the claim is mistaken, obviously, but it's not a trivially false and thus uninteresting horn of the trilemma.

2 See, for example, Franks & Wilson (2010), Kinna (2012) and McLaughlin (2007). For overviews of historical anarchist thought and theorists, see Guerin (2005) and Woodcock (1977), and for a strong argument against my claim that autonomy is (or ought to be) important for anarchism, see Bookchin (1995). For what it's worth, Bookchin seems to me to be taking autonomy to be exactly the kind of individualistic self-sufficiency we have already rejected; his notion of freedom is much closer to the relational account of autonomy endorsed in this thesis.

self-sufficiency which have been levelled at those who appeal to autonomy, for there is no reason to think that power and authority over one's life require complete freedom from obligations, or that one must be implausibly self-creating or rational to be powerful and authoritative in the relevant ways. Regarding authority, I posed the question of what would be lacking in a 'political state of nature' (to use Lane's framing) where we had no concept of authority. Primarily, I argued, the notion of political authority serves as a means of giving and justifying the normative claims of social institutions. To resolve co-ordination problems, and - more positively - to generate social goods which would be unattainable without co-operation, we require some means to indicate institutions and policies which reliably fulfil these goals. Thus, the conception of authority, as well as its success conditions, will be substantively informed: what it is to *be* an authority is different across different political theories, so that a liberal capitalist will require a different conception of authority from a socialist or a monarchist.

In Chapters 3 and 4, I argued for a social-relational account of autonomy which draws heavily from Oshana's and explored some of its implications.³ The social-relational view of autonomy holds that an agent must meet certain psychological preconditions, or what Friedman called competencies, such as a sense of self and an ability to make and follow through on plans. Meeting these preconditions does not, however, mean that the agent is autonomous: autonomy is a matter of standing in certain social relations which make one powerful and authoritative over one's life.

The notion of autonomy as relationally *constituted*, rather than merely being causally affected by social relations, is crucial: if an agent stands in (to use the standard example) relations of slavery, then she lacks power and authority over her life regardless of whether she authentically endorses such relations. If we are to understand autonomy as something which is valuable because it evinces respect for the idea of being able to shape how one's life goes, then we cannot say that the agent who is unable to perform such shaping is autonomous - even if this inability is something that they endorse. Conversely, being relationally

3 See Oshana (2006).

positioned such that you are powerful and authoritative means that you are autonomous, even if you'd rather not be. As in Oshana's example of the 'would-be surrendered woman', the relational account suggests - correctly - that the autonomous agent can be unhappy with their life and can even be resistant to being powerful over it. But autonomy is not a matter of getting everything you want, or of being entirely content with how your life is going: it is a matter of being positioned such that your wanting to do something is - again, *prima facie* - a reason to do it. Relations which constitute our being powerful and authoritative over our lives are those of non-domination: power is not to be understood as relentless desire-satisfaction or mastery of the universe, but as relational standing such that we are not vulnerable to arbitrary interference in our lives. We are authoritative, meanwhile, insofar as this relational standing is the result of our being granted recognition respect as agents - the external recognition, rather than internal belief, that our being an agent and wanting to *x* is a *prima facie* reason for us to *x* and for others to help us to *x*.

There are two implications of the relational thesis which are especially worth noting. The first is that this account of autonomy comes with an equality condition 'built in', for if you value autonomy you are required to value everybody else's autonomy equally. Although this can be interpreted in a vaguely Kantian fashion as meaning that you reason inconsistently if you value your own autonomy but nobody else's (an interpretation which, as it happens, strikes me as correct), that is not the interpretation which generates the equality condition here. Rather, we are committed to what I have called maximal equal autonomy - the set of relations which grant each agent as much power over their lives as is consistent with other agents having the same amount of power over *their* lives - because determining agents' relational standing according to morally arbitrary distinctions is, ultimately, always a mug's game. Even if I think that my self-direction is important and everyone else can go hang, it is irrational for me to support social relations which enable or deny powers of self-direction in an inegalitarian fashion; even if I am (as it happens) situated such that I am not immediately threatened by such inegalitarian relations, this is more or less a matter of luck, and we do not want our autonomy to rest on being lucky. In other words, this account of autonomy does not need a separate

account for why everyone should be granted the same ‘space’ for autonomy as ourselves: my account builds in an equality condition insofar as we are committed to equal maximal autonomy if we value autonomy at all (for nakedly self-interested reasons, if for no others), and need not reach for other arguments about, say, justice. The second implication, which I explored particularly in Chapter 4, is that the personal is political. That is to say, there are no principled distinctions which can be drawn between paradigm cases of ‘personal’ and ‘political’: on the relational account, changes which make an agent more (or less) powerful and authoritative over their life have the same effect regardless of whether they are the result of direct democracy, dictatorial fiat, or informal power hierarchies.

In Chapter 5, I returned to the problem of constructing a concept of political authority which is consistent with its being justified through considerations of autonomy. According to the functional analysis adopted in Chapter 2, the role of political authority is to enable or facilitate the production of co-operative goods as well as resolve co-ordination problems. What these goods are will, as suggested, be dependent on one’s pre-existing political convictions; for the purpose of the thesis, I argued that we want political authority primarily to promote and defend relations of maximal equal autonomy. I left aside the pluralist question: on my view, goods like security and shelter will be required for autonomy in any case, and the conception of autonomy in play entails that authorities should also serve justice and equality as necessary parts of promoting autonomy. Pluralists about value are thus to some extent accommodated anyhow, but given that my functional analysis of authority suggests a strongly substantive account of authority (in that my account will say that an authority’s legitimacy is in large part to do with getting the “right” result rather than simply following the correct procedures), it should be no surprise nor weakness that an explicitly autonomy-focussed concept of authority may not serve as an equally good concept of authority for other values.

Political authority being delimited by its ability to produce maximal equal autonomy gives us three desiderata for the concept: (1) political authorities must be able to produce morally binding obligations, (2) which are binding at

least in part as a result of the authority's institutional structure, and (3) these obligations must themselves be consistent with maximal equal autonomy.

These three desiderata, I argued in Chapter 6, are best met by what I called the autonomy-justified conception of authority. On this view, an institution is authoritative (or a directive it issues is authoritative) and thus generates political obligation if it meets two conditions. The Dependence Condition requires that the institution act or effect action in ways which reflect our reasons, including autonomy as a presumptive good: no institution may thus claim authority, however it is constructed and however many people vote for its proposals, if those proposals are autonomy-hostile. Similarly, authorities may not coerce us with regards to non-presumptive goods or reasons which we do not in fact have - the proposals must either reflect actual shared reasons, or reasons to pursue presumptive goods, in order to be authoritative. The Equality Condition, meanwhile, stipulates that institutions and proposals must act or effect action in ways which tend towards maximal equal autonomy. This condition serves to filter potential directives which would be friendly towards autonomy in theory but hostile in practice, and to help us select between exclusive autonomy-friendly proposals. If two proposals would both be consistent with agents being powerful and authoritative over their lives, but under conditions of moderate scarcity one of them is less accessible than another, or would require disproportionate resource allocation, then the Equality Condition tells us to prefer the most accessible (or least expensive, etc.) proposal. These conditions are individually necessary and jointly sufficient for authority and political obligation: an institution which meets both conditions can legitimately issue coercive directives, and legitimately justify this coercion on the basis of autonomy.

2. Autonomy-Justified Authority, the State, and Anarchy

There is now a coherent and unified picture of how autonomy and authority interact, and how the latter must operate in order to be justified by the former. I will finish by explicitly laying out why this picture strongly implies both

philosophical and political anarchism, without crudely lumping together all states as an indistinguishable and equivalent mass of 'bad'. First, I show that the account of autonomy I have given requires relational standing and recognition respect which it is impossible to attain in (or at least because of) the state. Then, I move from this claim to the argument that states cannot therefore claim autonomy-justified authority: there may potentially be institutions within the state which can, but the basic assumption that we have any *prima facie* duties to obey the law in virtue of the state being a legitimate political authority is incorrect.

2.1 Relational Autonomy and the State

Recall that, on my view of autonomy and authority, we are not maximally autonomous if we lack respect *in virtue of our agency*; and if institutions do not grant and defend recognition respect to their participants, then they lack authority. This gives us a very straightforward and, I think, fairly convincing reason to think that states are inconsistent with agents being maximally autonomous: they do never, and can never, respect agents *qua* agents.

At the absolute utmost, states grant respect to agents as state citizens, and often not even that. In order to be permitted to stand in a position of even basic respect within the various social-relational structures of, for example, the UK, one must be a British citizen (or one must be a very wealthy foreign citizen). Being an agent is not enough for one's wishes to have *prima facie* normative force under these conditions; if an agent is an asylum seeker or so-called illegal immigrant, then their wishes are not granted even secondary importance by the British government. Let us be clear about this: reflexively disregarding (as opposed to denying after consideration) somebody's desires is bad enough even if those desires concern cake, choice of film, or favoured seat. But refusing to take somebody's need for safety or security as a fact of the first importance is essentially to regard their *life* as unimportant. Sometimes, perhaps, it might transpire that these needs are already met, or that greater needs must be met first, but you *cannot* claim to recognise and respect an agent *qua* agent if their request for vital assistance doesn't stop you in your tracks. States, when they are concerned with agents at all, require a passport - or a skin - of the right

colour before they recognise that those agents' wishes might have some normative weights. My having been born in such-and-such a place is not a relevant factor in whether or not I need to eat and keep a roof over my head; but if I and a non-national go to the dole office, it is not agency nor need but citizenship which will determine who gets social security.⁴

These are, crucially, not contingent unpleasantnesses or mistakes in the functioning of a well-ordered society; these are *conceptual necessities of the statist hierarchy*.⁵ Any society which held the principle of recognition respect dear would not tolerate arbitrary divisions based on map lines, historical accident and injustice or superficial sexual characteristics. State organisation requires that there be those excluded from recognition, however, because without the power to exclude agents it is plausibly not a state. Weber's definition of the state as a "human community which successfully claims the monopoly on the legitimate use of force within a given territory" (2004, p33) has maintained its usefulness for a reason, after all. Here's an example: if nation-states cannot treat agents differently in virtue of, say, different nationalities, then no state has the right to eject or arrest non-citizens just because they are non-citizens. But this means that their claim to sole legitimate use of force within a territory is unsuccessful: they do not, in fact, have the right (let alone sole right) to use force within a given territory if they cannot eject people from it. This is why the common response that (some) states are bad, but (no) states *need* to be will not fly. States *must* base their limited recognition of agents on morally arbitrary factors like place of birth; indeed, they are required to do so

4 An inequality which the government are doing their best to erase via the simple expedient of making sure nobody gets any social security, but there we are.

5 It could be objected here that I'm deploying a concept-first analysis of the state for no good reason - why, the interlocutor can inquire, not just adopt another functional analysis such that there is no such conceptual necessity? This objection overlaps to some extent with the "state by any other name" criticism I consider later in the chapter, but is worth addressing here as well. Briefly, my response is this: neither the concepts of autonomy nor authority that I have used are made unrecognisable through being arrived at by functional analysis - autonomy is still to do with self-direction, and authority is still to do with legitimate coercion. But if we arrive at a concept of the state (as something like "the legitimate organising authority of some area or community", I take it) which features no powers of exclusion or sole legitimate use of force, I think we have given a concept which no longer bears *any* relation to its usual understanding. This doesn't, as it happens, strike me as a knock-down problem - as I suggest later on, if you want to call my picture of the anarchist society a state you're free to do so - but it does invite the question of why we should use that term rather than any other.

on pain of losing their main claim to statehood. And, to reiterate the point, if you deny somebody the things they need for survival or flourishing because of morally arbitrary factors, you cannot pretend to be granting them recognition respect as an agent.

The state is therefore incapable of providing one key requirement of being non-dominated and thus autonomous, namely recognition respect in virtue of agency. As I now recapitulate, non-vulnerability is also necessarily absent from the relations citizens find themselves in under statist organisations.

As I argued in Chapters 3 and 4, in order for us to be non-dominated, we must stand in relations which are hostile to arbitrary interference in our lives. If I am embedded in relations which mean that some person or group can decide how my life goes without reference to my own values, or the requirements of equal maximal autonomy, then I am vulnerable and hence dominated. The question is thus whether state organisation is compatible with relations of non-vulnerability, and the answer is that it is not.

To give a particularly relevant example: the institution of the police must, if it is to serve its primary function of maintaining the state and its laws, be in a position of domination over citizens. If we genuinely had protection against arbitrary interference in our lives, then the police could not be used as a method to suppress political and economic opposition. But this is, in fact, how the police are used: the threat or reality of police interference in one's life is a common and well commented upon tactic used to intimidate and silence challengers to particular policies or governments in general, and without the police having the ability to dominate citizens this tactic would not be effective.⁶

It's worth being clear about how this follows from what I've said about relational autonomy and the state: to focus narrowly on one aspect, agents A and B do not dominate each other so long as violence towards each other is equally forbidden (or sanctioned). That is, if the response to A clubbing B without provocation

⁶ It will not go unnoticed that this analysis both explains and has depressing implications regarding the currently high-profile murders of young black men by police in the United States (and, indeed, the British police's murders of Mark Duggan, Paul Tomlinson, Sheku Bayou and many others) and the probability of any effective change being made by governments.

would be the same as the response to B clubbing A without provocation, neither are in this respect dominant over the other. This means, of course, that A cannot prevent B from organising strikes which threaten a particular economic orthodoxy by beating them up and being free to do it again - and A is thus powerless to maintain the state's sole claim on legitimate violence. A is, in short, not a very useful police officer from the state's perspective. It will not do, either, to claim that since we elect representatives we have any meaningful ability to change this state of affairs through reform. As I have argued consistently throughout the thesis, one of the features of a hierarchical political and economic system is that the non-elite are significantly less powerful to change policy than the elite even if there is some formal equality: we thus face a situation where the legislative tools to render the police non-dominating are kept out of our hands, in large part thanks to the ability of police to arbitrarily interfere (or threaten to interfere) with our lives.⁷

Though there are plausible reasons for some kind of peace-keeping group in an anarchist society - appealing as, for example, Kropotkin's (1972) faith in humanity's basically benign nature is, it strikes me as unreliable and probably unjustified - such a group would look very different from the institution which exists first and foremost to protect the powerful and the current order of things.⁸

The relational account of autonomy thus entails the incompatibility of the state and maximal autonomy in at least two key ways. First, states do not and cannot grant recognition respect in virtue of our agency. Most obviously, the existence of closed borders show that, at most, states are prepared to grant limited recognition respect only to their citizens or non-citizens who pass a suitably stringent test. Second, there are some kinds of power structures entailed by

7 Of course, people can and do effect change through direct action and other means which do not involve taking the parliamentary path. The point is thus not that a suitably reformed parliamentary system would solve these problems, but that *no* parliamentary system can solve these problems and the police are key players in making direct action as difficult and dispiriting as possible. The game is rigged, in other words, and the machinery of the state is firmly behind making it the only game in town.

8 I specify "first and foremost" because the police clearly *do* play a role of protecting citizens, at least at some junctures. But this doesn't give us any reason to think that's their primary role, and nor does the fact that many or most of us *would* call the cops if we saw a gang beating up some victim: by analogy, I'd rather call Batman than let the Joker poison Gotham City, but given the option I'd rather be shot of them both.

what states *are*, and these structures are inconsistent with our being non-vulnerable: a society structured such that there is a set of institutions which citizens are dominated by is not a society consistent with maximal equal autonomy, and this is how state societies are and must be structured. Just as you don't get rich in an egalitarian society precisely *because* you don't get poor, and racism plays no functional role when white skin gives you no unfair advantage, a society structured for maximal autonomy would not and could not co-exist with an institution which exercised domination over members of that society. The police are one such institution, but so are states themselves: the former are simply one particular manifestation of the latter's dominance.

This is not, of course, to regurgitate the leftist fallacy that come the revolution, everything will be fine; you cannot just topple the superstructure and assume that the base will change, to borrow (or perhaps expropriate) the Marxist notion. Equally, it *is* possible, as I have indicated throughout this thesis, for some states to be better than others - there are obvious reasons to prefer liberal social democracy to a Victorian *laissez-faire* state, and to prefer either to Nazi Germany. Autonomy is a matter of the social relations that we actually find ourselves in, and those relations can be more or less constitutive of autonomy depending on what defines and controls the nature and interaction of those relations. Ultimately, though, if we desire a society where everyone stands in relations of power and authority over their lives, without domination of others, then we cannot be statist; and we ought, I think, to desire that society.

2.2 *Autonomy-Justified Authority and the State*

If we accept these arguments, then there is one trivially obvious reason why states may not claim autonomy-justified authority. Since maximal equal autonomy is inconsistent with statism, and a requirement of the autonomy-justified conception is that authorities claiming it act in ways consistent with maximal equal autonomy, states cannot claim autonomy-justified authority. It is still worth taking a moment to explore whether states can act to promote autonomy and in ways consistent with equal autonomy: I have already claimed that we can have autonomy-minded reasons to think of some state directives as

being authoritative, after all, and this apparent contradiction requires explanation.

I argue that while we do have reasons to obey or participate in at least some state directives, and that at least some of these reasons are given to us as a result of valuing autonomy, they do not amount to making the state authoritative (in the autonomy-justified sense). That is, we have autonomy-minded reasons to obey some of the state's laws, but no reasons which would line up with the state having autonomy-justified authority - there is something like political obligation which is given to us in virtue of our political goals, not in virtue of the putative authority being constituted such that it meets the Dependence and Equality Conditions. Thus, either the *state* cannot claim autonomy-justified authority but some of its constituent institutions can - which is in fact consistent with what I have already said about autonomy-justified authority - or we have fairly weak *prima facie* obligations to obey the state which are not actually political obligations (as in the case where we have an obligation to obey Nazi speed limits without this being a political obligation generated by the state). These options are not, I think, exclusive.

Take one example from the previous chapter, the introduction of equal marriage in the UK. I said there that, *pace* arguments about the institution of marriage, we seem likely to have equality-minded obligations to obey and work within laws extending the marriage franchise. The Equality Condition, I claimed, plausibly was met by a directive which so extended marriage, if we presumed that "more equality within the state" was better for autonomy than "less equality within the state", and that a directive legislating equal marriage would in fact make citizens more equal. Unless we are to believe that any state sanctioning (or expansion) is inconsistent with folks being more autonomous than before, then more equality within the state being better for autonomy, and the extension of marriage making folks more equal, seem like reasonable presuppositions.

Now, there *are* plausible worries about the way the decision has been taken in regards to the Equality Condition, in particular worries related to representative democracy. Even if we have the right result (in terms of the Equality Condition), it's been arrived at through a decision-making process which systematically

disempowers agents who aren't members of the economic or political boss class. Without dismissing such worries, though, I think we can still argue that they don't entail an inability to meet the Equality Condition. Since every state decision will have this feature of limited democratic legitimacy, it is not a question of the democratic shortfall requiring disobedience rather than compliance with this particular law, but of whether maximal equal autonomy is better served by compliance or disobedience. In this case, however, note that nothing about this qualified meeting of the Equality Condition suggests that we are obliged to obey the law and do nothing else. We are not, in other words, obliged to obey the law and refrain from trying to abolish marriage (or the state) entirely; we are merely obliged not to take up a contrarian position of disobeying any and all laws *because* they are the directives of a fundamentally undemocratic state.

So far, we either have a very tentative satisfaction of the Equality Condition, or simply a reason given by autonomy's internal equality requirement, which demands obedience to equal marriage laws. Given what I have said about the autonomy-justified conception previously, something's meeting the Equality Condition suggests that it should meet the Dependence Condition as well. If the proposal will in fact tend towards maximal equal autonomy, it seems likely (though not logically entailed) that it reflects the autonomy-minded reasons we have for obedience to authorities in the first place. Interestingly, though, that does not seem to be the case in respect of equal marriage laws. For example, our presumptive reasons regarding autonomy do *not* suggest that our cultural and relationship practices ought to be determined by representative democracy, and it is unlikely that every citizen actually has a non-presumptive reason to favour equal marriage.

We might, nonetheless, think that there is a move to be made which roughly parallels our reasons to think the Equality Condition still applies in non-ideal circumstances. We have no autonomy-minded reasons to bring about the state and its sanctioning of relationships, the argument runs, but given that we have both the state and state-sanctioned relationships, we have autonomy-minded reasons to expand the latter as broadly as possible. Given that autonomy is a

presumptive good, it is therefore not required for every citizen to have non-presumptive reasons to favour equal marriage, and the Dependence Condition is satisfied. As with the Equality Condition, any obligation as a result of this is very limited - we ought to obey the law, but have no requirement to then try and maintain the legal system.⁹

Now, though, there is an anarchist objection to field. After all this huffing and puffing, have I not ended up with an account of authority and anarchy which in the final analysis is pretty much indistinguishable from some kind of woolly reformism? If states can generate political obligation in virtue of being not quite so bad as they might be, then the Hobbesian or pseudo-Hobbesian claim that it is vital to maintain the state - because the alternative is worse - seems to follow along pretty closely behind.

I can reject this argument for two reasons. First, our obligations to obey “better than the current alternatives” laws are, as I have suggested, not obligations to obey them and then stop fighting for a better society. Certainly, we should be wary of having our organisational energies diverted into reformism; but nothing about the suggestion that we ought to obey equal marriage laws implies that we have to either aim to bring such laws about, or employ those tactics endorsed by liberal democracy to do so. Indeed, many laws which we can think of as better than nothing or better than the current alternatives - working time directives, employment rights, social security - have been brought about as the result of a government running scared in the face of direct action with much more fundamental goals than reform. That the laws exist doesn't mean that the work is done, but it seems plausible to say that we ought not to seek to overturn them just to prove how much more radical than thou we are or that we cannot be bought off with reformist sops.

Second, it still does not seem that we *do* have a general obligation to obey the state: if equal marriage laws, to stay with the same example, do in fact meet

⁹ Although this does, rightly, suggest that autonomy-motivated anarchism cannot just aim at removing the state. If we object to letting the state have the power to privilege some relationships over others, we are committed to changing society such that legal recognition is not required to ameliorate other oppressive structures, not just paring the powers of the state. The state is a particularly obvious and formal-institutional example of an unjust power hierarchy, but it is not the only one.

the Dependence and Equality Conditions, it won't be the case that every law does and thus generates obligations. If there is some challenge to equal marriage, for example, and a law is passed repealing the decision, then *ceteris paribus* there is no obligation to obey it. Under conditions of background inequality and democratic shortfall, in other words, some policy having been made an institutional directive is no reason at all to think that it generates obligations. If two directives pass exactly the same procedural requirements but have different authoritative status, then they are not getting their authority from the legislating institution but from external considerations (about autonomy, justice, etc.). To say that we might have some obligations to obey some laws, in short, is in no way to accept that the state's creation of a law gives us even a *prima facie* reason to obey it.

The answer to the question of whether states can claim autonomy-justified authority, then, is a negative but a complex one. States cannot claim autonomy-justified authority, *but* it is possible that some of their participant institutions can, iff it makes sense to think of those institutions as claiming political authority at all. It is also possible that states cannot claim autonomy-justified authority *but* may still issue directives which we are obliged to obey thanks to independent moral reasons.

It's conceivable, for example, that some hypothetical allotment committee set up by local government might be able to claim the autonomy justification despite being part of the machinery of the state, so that it can issue (very limited) authoritative demands. Equally, although we have good reasons to do what, for example, firefighters and paramedics tell us, it seems odd to say that they are members of institutions which have *political* authority - rather, they make morally obligatory demands which, as it happens, are co-extensive in this instance with the demands made by the state. In either case, that we have reasons to comply with some state institutions does not entail that the state has autonomy-justified authority.

2.3 *Autonomy-Justified Authority and Anarchy*

The objection above, that my account of authority ends up committing us to a liberal reformism in effect if not in theory, has a corollary which is worth addressing in its own right. The corollary is this: even if I'm correct that autonomy-justified authority does not explicitly require statism or reformism, any institution which can legitimately claim it will end up looking very like or in effect being a state. The "state by any other name" objection is familiar, but I rehearse it here for the sake of clarity.¹⁰

The criticism runs as follows. On my account, there is some institution (or more likely group of institutions) which can legitimately tell us what to do in some contexts. Further, these institutions do have powers of coercion - what powers exactly cannot be reasoned out *a priori*, but they must include the power and right to restrain and constrain folk under some circumstances, and to expropriate certain resources. If they don't have these powers then they cannot serve the role of defending and promoting relations of maximal equal autonomy, and so fail to be effective authorities. But powers of legitimate coercion over non-consenting individuals are surely, as Weber (*op. cit.*) suggests, the hallmark of the state. Now, perhaps the state implied by my account is a very different state to the ones we are familiar with - it is radically egalitarian, directly democratic, and very decentralised - but it is a state nonetheless. And an anarchism which includes a theory of the legitimate state is a very odd anarchism indeed.

One possible response here is just to dismiss the objection as a taxonomical quibble, and I don't think this response is entirely without merit. The anarchist objection to the state, after all, is not an objection to the idea that others may have legitimate authority over us, nor an objection to the existence of organising bodies.¹¹ If somebody wants to insist that any kind of organisational

¹⁰ See, for example, Miller's (1984) critical overview of anarchist thought.

¹¹ See, e.g., Malatesta for both an endorsement *and* a rejection of my historical claim here. On the one hand, he makes it clear that persistent non-compliers should be ejected from free associations: "these malcontents cannot fairly demand that the wishes of many others should be sacrificed for their sake" (1965, p29). This looks like a pretty clear endorsement of the idea that there can be organising bodies with legitimate rights to use force at least so far as necessary to exclude agents. On the other hand, he speaks in *Anarchy* of not using

network or set of institutions which can legitimately make demands is a state, it is tempting to shrug and let them have at it. But there is, I think, a key distinction between the state and the kinds of organisation that I argue can be legitimately authoritative, and a distinction which makes the taxonomical response redundant.

The distinction between anarchistic autonomy-justified authorities and the authority of states is that autonomy-justified authorities need claim neither sole nor general rights of coercion, while states claim both. Here is one example: the co-operative which oversees food production and distribution in my region determines that some area of fallow land should be turned over to growing something or other. We vote on it and approve the proposal - and check with other organising bodies to make sure that there are no competing proposals which don't involve food - but some die-hard grass-counter takes to making a nuisance of themselves so that planting can't begin. The committee will, I take it, endorse coercion to prevent this from happening, and will (by hypothesis) be justified in doing so. But there is nothing saying that forcibly removing the dissenter from the field has to be done by some agricultural compliance team, or that the removal must be cleared with them first. Similarly, there is no implication that the decision legitimates any coercion beyond removing the dissenter; the committee does not and cannot claim the right to, for example, exclude the dissenter from the distribution of resources as punishment.

Contrast this with state coercion: if a parallel decision is made, and the grass-counter again obstructs the carrying out of that decision, then the coercion is justified solely by reference to the authority of the state. There is a superficial similarity with the anarchist example - in both cases, a decision results in legitimate coercion - but when the food co-op's decision justifies coercion, it is the fact that the decision is in the interests of maximal equal autonomy which justifies the coercion. *That* they have made the decision is or should be a good

“expressions such as ‘abolition of the state’....substituting for it the clearer and more concrete term ‘abolition of government’” (p14), and of anarchy as “society without government” (p15) where more commonly anarchists speak of society without *rulers*. I tend to think that from what he says elsewhere in *Anarchy* and his other works that he is in fact using “government” to refer to a representative or hierarchical system of rulership, rather than simply an institution for co-ordinating and producing public goods, but the reader may disagree.

sign that the proposal does indeed defend or promote autonomous relations, but it is not the case that coercion is justified simply because the committee has made the decision. Our right to coerce non-compliers in this case is suggested by, not dependent upon, the committee decision; when the state sanctions coercion, any rights to coerce are solely derivative from their being sanctioned by the state. A group of people who removed the agent from the land without state approval might very well find themselves up on charges.

So, states claim sole right of coercion where autonomy-justified authorities do not; they also claim general rights of coercion where autonomy-justified authorities do not. State coercion of the grass-counter will not be limited to their removal from this particular piece of land. Non-compliance with a particular state directive results in a general forfeiture of rights against coercion. The obvious instance would be when, in our example, the state decides to imprison the dissenter and justifies this very broad coercion by reference to its own authority. That is to say, having claimed the right to coerce the individual just by dint of that individual's disobedience of the law, the state then claims the right to coerce and punish them in ways which are unconnected with and do not further the original proposal. The coercion is a general punishment for lawbreaking, not a means to ensure the efficacy of the food production directive.

3. Concluding Remarks

In Section 2, I reran some anarchist arguments against the state and police in the context of my accounts of autonomy and authority, and argued that an anarchist society with autonomy-justified authorities was neither self-contradictory nor a state by another name. What I have not done is give any concrete account of how we are to move from our hierarchical, statist society to the anarchism of social-relational autonomy I have advocated in the thesis. Partly this is down to constraints of space, but mostly it's because I have no exhaustively worked-out revolutionary how-to guide.

Still, I think some rough suggestions are implied by the preceding theoretical fandango. First, the anarchist core principle of bottom-up organisation is reaffirmed by the account of autonomy as relationally constituted and requiring non-domination. A government policy or reform which (say) increases social security may very well make folk better off than they were before, but the structure of the reform - proceeding from party lines or calculations of political expediency, and imposed by a dominating body - is still analogous to slaveholders deciding to make their slaves more comfortable in order to get more work out of them or stave off a revolt. When we organise directly, however, shutting down factories, resisting evictions, and so on, we are (albeit incompletely and perhaps temporarily) making ourselves less vulnerable to interference from those dominating agents by restructuring the relations ourselves in a way which is both consistent with, and required for, greater equal autonomy for agents.

Second, I have shown that the charge of hyper-individualism against anarchism, in particular autonomy-motivated anarchism, does not stick. The argument which does this, however - the argument that relational autonomy requires relations of equal autonomy - entails that we have obligations of solidarity and support which are far from limited to resisting the state. The state sanctions and enforces racism, misogyny, homophobia, inequality; but so too do non-state organisations. When we reject the police as a means of fighting fascism and racism, we must do it ourselves, organising directly against fascist and racist groups and working to eradicate hierarchies of privilege and power. If autonomy-minded anarchism is to be consistent with a social rather than individualist outlook, in other words, it must consist of more than anti-statism.

Third, and relatedly, to engage with fundamentally and fatally flawed institutions is not to be a bad anarchist, or a dupe of the state - at least, not necessarily. My membership of a reformist trade union is indeed predicated on the existence of capitalism; reformist trades unions essentially exist to mediate the relationship between boss and worker, and that is not a relationship which exists in an egalitarian and anarchist society. But it is still a tool - a limited, ultimately unreliable tool - to make things a little bit better. We do badly as

anarchists if we rely on union bureaucracy and government mediation to bring about a better society, but I think we also do badly to insist on doctrinal purity over material improvements in our relational standing. Reform at the cost of revolution would not be worth fighting for; but it is very rare that we actually find ourselves faced with such a dichotomy, and it does nobody any favours to dismiss concrete victories as mere chimerae. You can't, to quote Howard Zinn (and a thousand lefty think-pieces), be neutral on a moving train, and the folk fighting for higher wages or better healthcare are not our class enemies just in virtue of their statism.

Ultimately, and perhaps appropriately, my thesis' novelty lies more in the way it explains the relation between autonomy, authority and anarchy than in any particular definition or description of the concepts. If we value self-direction and inter-personal relationships, we ought to adopt relational autonomy; if we are relationalists about autonomy, then we ought to be anarchists; and if we are social rather than individualist anarchists, then we ought to have some working concept of political authority. This extended and tortuous formulation is not, I admit, anywhere near as catchy as "I am not free until everybody around me is free" - but it amounts to saying the same thing.

Bibliography

- Alford, J., and Hibblings, J., 'Personal, Interpersonal and Political Temperaments', *Annals of the American Academy of Political and Social Science* (Vol. 614, 2007), pp196-212.
- Arendt, H., *On Violence* (Penguin: London, 1970).
- Bakunin, M. (ed. Shatz, M.), *Statism and Anarchy* (Cambridge University Press: Cambridge, 2005).
- Bartky, S.L., *Femininity and Domination: Studies in the Phenomenology of Oppression* (Routledge: New York, 1990).
- Beatty, C. and Fothergill, S., *Hitting the poorest places hardest: The local and regional impact of welfare reform* (Centre for Regional Economic and Social Research: Sheffield Hallam University, 2013).
- Benn, S.I., 'Freedom, Autonomy and the Concept of a Person', *Proceedings of the Aristotelian Society* (Vol. 76, 1975-76), pp109-130.
- Berlin, I., *Four Essays on Liberty* (Oxford University Press: London, 1969).
- Bookchin, M., *Social Anarchism or Lifestyle Anarchism* (AK Press: Edinburgh, 1995).
- Brake, E. (ed.), *After Marriage: Rethinking Marital Relationships* (Oxford University Press: New York, 2016).
- Bratman, M., 'Autonomy and Hierarchy' in Sen and Williams (eds., 1982), pp156-176.
- Brennan, S., Isaac, T.L., and Milde, M. (eds.), *A Question of Values: New Canadian Perspectives in Ethics and Political Philosophy* (Radopi: Amsterdam, 1997).

Carr, C., 'The Problem of Political Authority', *The Monist* (Vol. 66, No. 4, 1983), pp472-486.

Carter, I., 'Respect and the Basis of Equality', *Ethics* (Vol. 121 No. 3, 2011), pp538-571.

Chambers, C., 'The Limitations of Contract: Regulating Personal Relationships in a Marriage-Free State' in Brake (ed., 2016), pp51-83.

Christiano, T., 'Democracy as Equality' (1993). Available online at <http://philosophyfaculty.ucsd.edu/faculty/rarneson/Courses/christianodemocracy.pdf> (accessed 24/05/13).

Christman, J., 'Constructing the Inner Citadel: Recent Work on the Concept of Autonomy', *Ethics* (Vol. 99 No. 1, 1988), pp109-124.

--- (ed.) *The Inner Citadel: Essays on Individual Autonomy* (Oxford University Press: New York, 1989).

--- 'Autonomy and Personal History', *Canadian Journal of Philosophy* (Vol. 21, No. 1, 1991), pp1-24.

--- 'Defending Historical Autonomy: A Reply to Professor Mele', *Canadian Journal of Philosophy* (Vol. 23, No. 2, 1993), pp281-289.

--- *The Politics of Persons* (Oxford University Press: New York, 2009).

Christman, J. and Anderson, J. (eds.), *Autonomy and the Challenges to Liberalism* (Cambridge University Press: Cambridge, 2005).

Code, L., 'The Perversion of Autonomy & the Subjection of Women' in Mackenzie and Stoljar (eds., 2000), pp181-209.

Cohen, G.A., *History, Labor and Freedom* (Oxford University Press: Oxford, 1989).

Colburn, B., *Autonomy and Liberalism* (Routledge: New York, 2010).

--- 'Autonomy and Adaptive Preferences', *Utilitas* (Vol. 23, 2011), pp52-71.

Cooke, M., 'Five Arguments for Deliberative Democracy', *Political Studies* (Vol. 18, No. 5, 2000), pp947-969.

Craig, E., *Knowledge in the State of Nature* (Oxford University Press: Oxford, 1999).

Darwall, S., *The Second-Person Standpoint: Morality, Respect and Accountability* (Harvard University Press: Massachusetts, 2009).

Dimock, S., 'Personal Autonomy, Freedom of Action, and Coercion' in *A Question of Values: New Canadian Perspectives in Ethics and Political Philosophy*, pp65-87.

Dworkin, G., 'Autonomy and Behavior Control', *The Hastings Center Report* (Vol. 6, No. 1, 1976), pp23-28.

--- *The Theory and Practice of Autonomy* (Cambridge University Press: Cambridge, 1988).

--- 'The Concept of Autonomy' in Christman (ed., 1989), pp54-62.

Dworkin, R., *Sovereign Virtue* (Harvard University Press: London, 2002).

Elster, J., 'Sour Grapes: Utilitarianism and the Genesis of Wants' in Sen and Williams (eds., 1982) pp219-238.

--- 'The Market and the Forum: Three Varieties of Political Theory' in Goodin and Pettit (eds., 1997) pp128-142.

Eyerman, R., 'False Consciousness and Ideology in Marxist Theory', *Acta Sociologica* (Vol. 81, No. 1/2, 1981), pp43-56.

Finlayson, L., *The Political is Political: Conformity and the Illusion of Dissent in Contemporary Political Philosophy* (Rowman and Littlefield International:

London, 2015).

Fischer, J.M., & Ravizza, M., 'Responsibility and Inevitability', *Ethics* (Vol. 101 No. 2, 1991), 258-278.

--- *Responsibility and Control: A Theory of Responsibility* (Cambridge University Press: Cambridge, 1998).

Frankfurt, H., 'Freedom of the Will and the Concept of a Person', *Journal of Philosophy* (Vol. 68, No. 1, 1971), pp5-20.

Franks, B. and Wilson, M. (eds.), *Anarchism and Moral Philosophy* (Palgrave Macmillan: Basingstoke, 2010).

Friedman, M., *Autonomy, Gender, Politics* (Oxford University Press: New York, 2003).

--- 'Autonomy and the Split-Level Self', *Southern Journal of Philosophy* (Vol. 24, No. 1, 1986), pp19-35.

Goodin, R. and Pettit, P. (eds.), *A Companion to Contemporary Political Philosophy* (Blackwell: Oxford, 1993).

Goodin, R. and Pettit, P. (eds.), *Contemporary Political Philosophy: An Anthology* (Blackwell: Oxford, 1997).

Gould, C., *Rethinking Democracy: Freedom and Social Co-operation in Politics, Economy and Society* (Cambridge University Press: New York, 1988).

Gramsci, A. (trans. Hoare, Q. and Nowell-Smith, G.), *Selections from the Prison Notebooks of Antonio Gramsci* (International: New York, 1971).

Guerin, D., *No Gods, No Masters* (AK Press: Edinburgh, 2005).

Gutmann, A., 'Democracy' in Goodin and Pettit (eds., 1993), pp411-420.

Hanrahan, R & Antony, L., 'Because I Said So: Towards a Feminist Theory of

Authority', *Hypatia* (Vol. 20, No. 4, 2005), pp59-79.

Hansson, S., 'Social Choice with Procedural Preferences', *Social Choice and Welfare* (Vol. 13, No. 2, 1996), pp215-230.

Harrison, R., *Democracy* (Routledge: London, 1993).

Hart, H.L.A., 'Are There Any Natural Rights?', *The Philosophical Review* (Vol. 64, No. 2, 1955), pp175-191.

Hill, T., 'Servility and Self-Respect', *The Monist* (Vol. 57, No. 1, 1973), pp87-104.

--- *Autonomy and Self-Respect* (Cambridge University Press: Cambridge, 1991).

Hobbes, T., *Leviathan* (1651. Text used: Project Gutenberg, <https://www.gutenberg.org/files/3207/3207-h/3207-h.htm>, accessed 18/11/2013).

Hume, D., *Of the Original Contract* (1748. Text used: Constitution Society, <http://www.constitution.org/dh/origcont.htm>, accessed 02/03/2016).

Humphries, J., 'Intimacy, Autonomy and (Non) Domination', *Journal of Applied Philosophy* (forthcoming).

Kaplan, L., 'Woman as Caretaker: An Archetype That Supports Patriarchal Militarism', *Hypatia* (Vol. 9, No. 2, 1994), pp123-133.

Kay, A., Jordan, H., and Baker, M., *The Tipping Point: The human and economic costs of cutting disabled people's support* (Hardest Hit coalition, 2012).

Kinna, R. (ed.), *The Bloomsbury Companion to Anarchism* (Continuum: London, 2012).

Klosko, G., 'The Principle of Fairness and Political Obligation', *Ethics* (Vol. 97, No. 2, 1987), pp353-362.

--- 'Political Obligation and the Natural Duties of Justice', *Philosophy and Public Affairs* (Vol. 23, No. 3, 1994), pp251-270.

--- *Political Obligations* (Oxford University Press: Cambridge, 2005).

Knowles, D. & Skorupski, J. (eds.), *Virtue & Taste* (Blackwell: Oxford, 1993).

Korsgaard, C., *Creating the Kingdom of Ends* (Cambridge University Press: Cambridge, 1996).

--- *Sources of Normativity* (Cambridge University Press: Cambridge, 1996).

Krakauer, J., *Into the Wild* (Villard Books: New York, 1996).

Kropotkin, P. (ed. Avrich, P.), *Mutual Aid: A Factor of Evolution* (Allen Lane: London, 1972).

Lane, M., 'States of Nature, Epistemic and Political', *Proceedings of the Aristotelian Society* (New Series, Volume 99, 1999), pp211-224.

Lee, T.M.L., 'Rethinking the Personal and the Political: Feminist Activism and Civic Engagement', *Hypatia* (Vol. 22, No. 4, 2007), pp163-179.

Levy, N., 'Culpable Ignorance and Moral Responsibility: A Reply to Fitzpatrick', *Ethics* (Vol. 119 No. 4, 2009), pp729-741.

Mackenzie, C. and Stoljar, N. (eds.), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency and the Social Self* (Oxford University Press: Oxford, 2000).

Malatesta, E. (trans. Richards, V.), 'A Talk About Anarchist Communism Between Two Workers', *Errico Malatesta: His Life and Ideas* (Freedom Press: London, 1965).

--- *Anarchy* (Freedom Press: London, 1974).

Marx, K. (trans. Quelch, H.), *The Poverty of Philosophy* (Cosimo, Inc: New York, 2008).

--- (ed. McLellan, D.), *Karl Marx: Selected Writings* (Oxford University Press: New York, 2010).

McLaughlin, P., *Anarchism and Authority* (Ashgate: Aldershot, 2007).

Meyers, D.T., *Self, Society, and Personal Choice* (Columbia University Press: New York, 1989).

--- 'Decentralizing Autonomy: Five faces of Selfhood' in Christman and Anderson (eds., 2005), pp27-55.

--- 'Review of *Personal Autonomy in Society*', *Hypatia* (Vol. 23 No. 4, 2008), pp202-206.

Mill, J.S., *On Liberty* (Penguin: London, 1974).

Miller, D., *Anarchism* (J.M. Dent & Sons: London, 1984).

Mokrosinksa, D., 'What is Political About Political Obligation?', *Critical Review of International Social and Political Philosophy* (Vol. 16, No. 1, 2012), pp88-108.

Moore, A. and Lloyd, D., *V for Vendetta* (Vertigo/DC Comics: London, 1982).

Moore, G.E. (ed. Baldwin, T.), *Principia Ethica* (Cambridge University Press: Cambridge, 1993 (revised edition with "Preface to the second edition" and other papers)).

Noggle, R., 'Review of *Personal Autonomy in Society*', *Journal of Value Inquiry* (Vol. 45 No. 2), pp233-238.

Nozick, R., *Anarchy, State and Utopia* (Basic Books: New York, 1974).

Oshana, M., 'Ascriptions of Responsibility', *American Philosophical Quarterly* (Vol. 34, No. 1, 1997), pp71-83.

--- 'The Misguided Marriage of Responsibility and Autonomy', *Journal of Ethics* (Vol. 6 No. 3, 2002), pp261-280.

--- *Personal Autonomy in Society* (Ashgate: Aldershot, 2006).

--- (ed.), *Personal Autonomy and Social Oppression: Philosophical Perspectives* (Routledge: New York, 2014).

Otsuka, M., 'Is the Personal Political? The boundary between the public and the private in the realm of distributive justice' (English version of a paper published in Italian), *Iride: Filosofia e Discussione Pubblica* (Vol. XIV, No. 34, December 2001), pp617-23.

Paul, E.F., Miller, F.D., and Paul, J. (eds.), *Autonomy* (Cambridge University Press: New York, 2003).

Pettit, P., 'Republican Freedom and Contestatory Democratization' in Shapiro, I. and Hacker-Cordon, C. (eds., 1993), pp163-190.

--- 'Freedom as Anti-Power', *Ethics* (Vol. 106, No. 3, 1996).

--- *Republicanism: A Theory of Freedom and Government* (Oxford University Press: New York, 1997).

Rand, A., *Philosophy: Who Needs It?* (Bobbs-Merrill Company: Indianapolis, 1982).

Rawls, J., *Political Liberalism* (Columbia University Press: New York, 1996).

--- *A Theory of Justice* (Harvard University Press: Massachusetts, 1999).

Raz, J., *The Authority of Law: Essays on Law and Morality* (Oxford University Press: Oxford, 1979).

--- *The Morality of Freedom* (Oxford University Press: New York, 1986).

Rousseau, J-J. (trans. Cranston, M.), *The Social Contract* (Penguin: Harmondsworth, 1968).

Schier, F., 'The Kantian Gulag', in Knowles and Skorupski (eds., 1993), pp1-9.

Schofield, P., *Utility & Democracy: The Political Thought of Jeremy Bentham*, (Oxford University Press: Oxford, 2006)

Sen, A and Williams, B. (eds.), *Utilitarianism and Beyond* (Cambridge University Press: Cambridge, 1982).

Shapiro, I. and Hacker-Cordon, C. (eds.), *Democracy's Value* (Cambridge University Press: Cambridge, 1999).

Simmons, A.J., 'The Principle of Fair Play', *Philosophy and Public Affairs* (Vol. 8, No. 4, 1979), pp307-337.

--- 'The Anarchist Position: A Reply to Klosko and Senor', *Philosophy and Public Affairs* (Vol. 16, No. 3, 1987), pp269-279.

Singer, P., *Democracy and Disobedience* (Oxford University Press: Oxford, 1974).

Steadman, J., 'Moral Responsibility and Motivational Mechanisms', *Ethical Theory and Moral Practice* (Vol. 15 No. 4, 2012), pp473-492.

Stirner, M. (ed. Leopold, D.), *The Ego and Its Own* (Cambridge University Press: Cambridge, 2009 (new edition)).

Stoljar, N., 'Autonomy and the Feminist Intuition' in Mackenzie and Stoljar (eds., 2000), pp94-111.

Thoreau, H.D., *Walden; or, Life in the Woods* (Dover Publications: New York, 1995).

Tomasi, J., 'Should Political Liberals Be Compassionate Conservatives?', *Social Philosophy and Policy* (Vol. 21, No. 1, 2004), pp322-345.

Wallace, J., 'Duties of Love', *Proceedings of the Aristotelian Society*, Supplementary Vol. LXXXVI (2012), pp175-198.

Weber, M., (trans. Livingston, R., eds. Owens, D. & Strong, T.). 'Politics as Vocation' in *The Vocation Lectures* (Hackett Books: Illinois, 2004).

Wedgwood, R., 'Is Civil Marriage Illiberal?' in Brake (ed., 2016), pp29-50.

Westlund, A., 'Selflessness and Responsibility for Self. Is Deference Compatible

With Autonomy?', *The Philosophical Review* (Vol. 112, No. 4, 2003), pp483-523.

Williams, J., 'The Personal is Political: Thinking Through the Clinton/Lewinsky/Starr Affair', *Political Science and Politics* (Vol. 34, No. 1, 2001), pp93-98.

Wolf, S., *Freedom Within Reason* (Oxford University Press: Oxford, 1994).

Wolff, R.P., *In Defense of Anarchism* (University of California Press: Oakland, 1998 (revised edition)).

Woolf, V., *Collected Essays* (Hogarth Press: London, 1966).