

On the Relational Conditions for Autonomy

By Jules Holroyd

Submitted for the degree of
Doctor of Philosophy
(Philosophy)

August 2008

The Department of Philosophy
The University of Sheffield

Abstract

This thesis is about the concept of personal autonomy. In particular, it is about how we might understand the claim that social relations are constitutive of autonomy, and whether, appropriately understood, it is true. Personal autonomy, broadly construed, can be understood as an individual's ability to govern herself: to interact with the world, deliberate about what to do, to choose, and execute her choice. We value autonomy not only because we value being self-governing, but also because the concept plays important roles in various theoretical frameworks: in grounding respect, in setting the bounds of legitimate paternalistic intervention, and in identifying the candidates for political participation. Recently, conceptions of autonomy that are 'relational' have been developed. Such conceptions make explicit reference to the social environment of the agent. This thesis assesses the prospects for an adequate relational conception of autonomy.

Acknowledgements

The writing of this thesis has been a challenging and enjoyable undertaking, made so by the supporting roles of many, to whom I am most grateful.

Many thanks to my supervisors, Jenny Saul and Jimmy Lenman, who have been a source of support and encouragement throughout. I have learnt much from discussion with them, not least that philosophy can be much fun. I am enormously grateful for their generosity in giving of their advice, time and support.

The friendly and enthusiastic environment in the department of philosophy at The University of Sheffield has been a wonderful setting to undertake a PhD, and I am grateful to all who have contributed to making it so. In particular, I have appreciated the opportunity to participate in the stimulating reading groups, in feminist philosophy and in ethics, for the past five years.

I am grateful to the following institutions, which provided significant financial support: the AHRC, for funding my PhD studies, and also for their generosity in supporting my overseas study visit; the Learned Society Fund of The University of Sheffield, who generously provided the funds that enabled me to attend a number of conferences. These experiences were of great value.

My time overseas at MIT in 2005 was productive and the progress I made there contributed significantly to the work of this thesis. For this, and for much encouragement, I am also indebted to Sally Haslanger and Rae Langton.

My particular thanks extend to the following good friends, with whom I have enjoyed much discussion philosophical and otherwise: Laura Beeby, Giles Banning-Lover, Fabian Freyenhagen, Leah Henderson, Gerry Hough, Anna Hughes, David Liggins, Anna Mahtani, Mari Mikkola, Joe Morrison, Lina Papadaki, Lindsey Porter, Yonatan Shemmer, Anna Wilkinson, Nick Wiltsher, and Rich Woodward.

Thanks to my family – Steve, Ginny, David and Linval Holroyd – who have provided unwavering support, throughout my PhD and beyond. Finally, thanks to Hadiru Mahdi, for inspiration.

Contents page

Abstract	i
Acknowledgements	ii
Contents page	iii
Introduction	1
Chapter 1: The Value of Autonomy	3
1.1 Why care about autonomy	3
1.2 Towards relational autonomy	9
1.3 Aspects of autonomy	14
1.4 A framework	16
Chapter 2: Substantive account I: Stoljar	19
2.1 Stoljar's 'feminist intuition'	21
2.2 Some content neutral conditions	23
2.3 Stoljar's substantive condition	35
2.4 Summary	43
Chapter 3: Substantive account II: Hill Jnr., Baron.	45
3.1 The problem with deference	46
3.2 Moral rights	50
3.3 Interests or preferences	53
3.4 Summary	62
Chapter 4: A schematic argument against substantive accounts.	64
4.1 A concern about values	65
4.2 Three ways to understand the substantive condition	68
4.3 Epistemic conditions for autonomous choice	71
4.4 World-guided conditions for autonomous choice	79
4.5 A hybrid condition	82
4.6 Problems with the fully specified condition	84
4.7 Summary	89
Chapter 5: Relational accounts of autonomous agency: Benson and Oshana.	93
5.1 Benson's account of autonomous agency	93
5.2 Oshana's account of autonomous agency	102
5.3 Substantive concerns	120
5.4 Summary	124
Chapter 6: An argument against constitutively relational conditions for autonomous agency.	126
6.1 Paternalism and autonomy	126
6.2 Relational conceptions and the core case	133
6.3 Three options	136
6.4 Conclusions	144
Chapter 7: Relational conditions for autonomous action.	146
7.1 More on autonomous action	147
7.2 Virtues of a relational conception of autonomous action	161
7.3 Summary	169
Conclusion	171
Bibliographic references	172

Introduction

This thesis is about the concept of personal autonomy. In particular, it is about how we might understand the claim that social relations are constitutive of autonomy, and whether, appropriately understood, it is true. Personal autonomy, broadly construed, can be understood as an individual's ability to govern herself: to interact with the world, deliberate about what to do, to choose, and execute her choice. We value autonomy not only because we value being self-governing, but also because the concept plays important roles in various theoretical frameworks: in grounding respect, in setting the bounds of legitimate paternalistic intervention, and in identifying the candidates for political participation. Recently, conceptions of autonomy that are 'relational' have been developed. Such conceptions make explicit reference to the social environment of the agent. It is with the prospects for an adequate relational conception that I shall be concerned in this thesis.

One aim of this thesis is to clarify the form of relational conditions, and hence gain a clearer view of the objections that such views may face. Can a relational conception adequately identify those individuals who are intuitively self-governing and to whom these normative benefits (of entitlement to respect, and so on) should be accorded? Do relational conceptions unavoidably incorporate values into the notion of autonomy? Does this mean that a relational conception cannot cohere with the normative frameworks that in part explain its significance? These are questions I shall address in this thesis. I will argue against existing theories of relational autonomy. However, I will ultimately argue that there are relational conditions for autonomous action.

I start on this project here by setting out some of the frameworks in which autonomy plays a central role, and which, in part, explain its value. This not only motivates the thought that we should care about autonomy, but also sets out the key roles in which a conception of autonomy should be able to function.

With a clearer sense of the concern, I turn, in the rest of the thesis, to look at some relational conceptions that have been formulated. It is useful, in doing so, to carve up the notion of ‘autonomy’ into its three aspects: autonomous choice, the standing as an autonomous agent, and autonomous action. Part 1 of the thesis considers and rejects a relational conception of autonomous choice. Part 2 assesses the success of relational conditions for being an autonomous agent, concluding that they should not be accepted. The conclusion of these first two parts serves to illustrate the difficulty of offering a relational conception that is able play the roles – protecting from paternalism, grounding respect, and so on – that explain autonomy’s value. The positive argument of the thesis comes in Part 3, where I set out the benefits of recognising autonomous action to be relationally constituted. I outline a framework for thinking about how social relations can thwart or enable the exercise of autonomy, and motivate the further pursuit of this line of thought.

Chapter 1. The Value of Autonomy

1.1. Why care about autonomy?

Some hold that autonomy has intrinsic value. But autonomy also plays a central role in important normative frameworks. In addition to its purported intrinsic value, we care about autonomy because being autonomous – and, crucially, being recognised as autonomous – comes with certain ‘normative benefits’. An autonomous agent is accorded a certain kind of respect; an autonomous agent is entitled to protection from paternalistic intervention; an autonomous agent is a candidate for political participation. These normative benefits not only account for why we care about autonomy; that autonomy has such a role makes clear just how important it is that an adequate conception of autonomy is formulated. An inadequate formulation may lead to diagnostic error, the practical import of which may be that some agents miss out on the normative goods to which they are entitled.

1.1.1. Intrinsic value

Commonsense morality (at least, of liberal western societies) esteems autonomy – self-governance is valued.¹ When an agent is prevented from governing herself – by another, or by circumstance, or by some deficit on her part – we express concern. Being self-governing is deemed to be valuable ‘in itself’; not because of the ends it serves, or the likelihood that it will best serve one’s interests.² As James Griffin puts it: “Even if I constantly made a mess of my life, even if you could do better if you took charge, I

¹ I say little here about different value systems that may repudiate the value of autonomy. However, I intend all my claims here to be compatible with value pluralism; that is to say, I do not take autonomy to be the sole value, nor even the most important. Rather, I take it that autonomy is important, and insofar as it is, we should want to consider carefully the ways in which social contexts might thwart or foster it.

² See Mill, J.S. *Utilitarianism*, Warnock, M. (ed.), *Mill: Utilitarianism and Other Writings*, (Collins, Glasgow, 1962) for the unlikely claim that our ends will best be served if we are left to make our own decisions: “with respect to his own feelings and circumstances, the most ordinary man or woman has means of knowledge immeasurably surpassing those that can be possessed by anyone else”, at p.207. This fails to take into account the enormous capacity for lack of self-knowledge, motivated irrationality, and general ignorance. On such a view, autonomy has merely instrumental value.

would not let you do it. Autonomy has a value of its own.”³ This thought is plausible; there is something objectionable about others determining one’s choices and actions, and there is something valuable about self-determination. Note, though the strength of the claim here, from Griffin: even if self-governance led to a much worse outcome, in terms of personal well-being, than were one’s autonomy to be overridden, it would be better to be self-governing than not. The claim, then, is not merely that autonomy is valuable, but that it is valuable to the extent that being self-governing might reasonably be chosen over avoiding ‘making a mess’ of things. Even if it thwarts our ends, there is value in being autonomous. Those who agree with Griffin may hold that autonomy has intrinsic value. This is not to say that autonomy can never be overridden, but simply that if it is, then justification is required – perhaps strong justification. If autonomy is intrinsically valuable, then we have at least prima facie reason to care about it.

For the purposes of this thesis, I will take it that autonomy is valuable. I find Griffin’s thought a plausible one. But it is certainly possible that some may not. This is why it is important to attend to the normative benefits to which autonomous agents are entitled: these further explain why autonomy is valuable and why we should care about it. If a conception of autonomy is ill-placed to play any such roles, then it becomes largely inert, robbed of its richness as a normative concept. We care about autonomy not only because we care about doing things for ourselves, but also because we care about what it means, normatively speaking, to be autonomous; namely, to be entitled to certain normative benefits.

1.1.2 Respect, citizenship

Two normative frameworks in which the notion of autonomy operates will not play a significant role in this thesis, but it is worth setting them out, briefly, here, in order to further explain the value of autonomy.

The first normative framework is that of grounding respect. One of the most prominent philosophers to connect autonomy to respect is Kant, who, with his Formula of Humanity, prescribed as a categorical imperative that each must respect humanity – the capacity for rational self-legislation – in oneself and in others.⁴ The relationship

³ Griffin, J. (1988) *Well-being: Its Meaning, Measurement and Moral Importance*, Oxford, Oxford University Press, p. 67

⁴ See Kant, I. *Groundwork of the Metaphysics of Morals* Gregor, M. (ed., transl). (1998) Cambridge, Cambridge University Press.

between Kantian moral autonomy and personal autonomy is disputed.⁵ But the thought that the distinctive capacity for reflective self-governance demands respect has persisted.

Stephen Darwall describes as ‘recognition respect’ the “sort of respect which is said to be owed to all persons”, in virtue of their standing as autonomous,⁶ contrasting this to the kind of respect that we might accord to the particular features or character traits of a person, manifested in certain pursuits – excellence in musicianship, or archery, say. John Christman too, talks of “autonomy [having]...value simply because it constitutes, in part, the human agency and capacity for authentic choice that grounds respect for ourselves and others as persons”.⁷ Likewise, we find the thought that there is an important value involved in making one’s own choices in T.M. Scanlon’s claim that:

in a situation in which people are normally expected to determine outcomes of a certain sort through their own choices unless they are not competent to do so, I may value having a choice because my not having it would reflect a judgment on my own or someone else’s part that I fell below the expected standard of competence.⁸

Because of the relation of autonomy to respect, making one’s own choice has a ‘symbolic’ value. Like Griffin, Scanlon holds that agents may want to make their own choices ‘for better or for worse’; this thought is in part explained by the claim that failing to choose for oneself involves failing to claim – or be regarded as having – status as a competent agent.

Autonomy, then, plays a role in grounding respect; though the arguments that attempt to establish this will not be the subject of this thesis.⁹ That autonomous agents

⁵ For discussion of the relationship between the two senses, see, for example, Gaus, G. (2005) ‘The Place of Autonomy within Liberalism’ *Autonomy and the Challenges to Liberalism* Christman, J. and Anderson, J. (eds.) 2005, Oxford, Oxford University Press, pp.272-306; O’Neill, O. (2003) ‘Autonomy: The Emperor’s New Clothes’ *Aristotelian Society Supplementary Volume* 77(1), pp.1-21.

⁶ Darwall, S. (1977) ‘Two Kinds of Respect’ *Ethics* 88 (1) pp.36-49. See p.38.

⁷ Christman, J. (2004) ‘Relational Autonomy, Liberal Individualism, and the Social Constitution of Selves’ *Philosophical Studies* 117, pp.143-164, see p. 153.

⁸ Scanlon, T.M. (1986) ‘The Significance of Choice’ *The Tanner Lectures on Human Values*, delivered Brasenose College, Oxford, May 1986. Lecture 2, at p. 181.

⁹ See e.g. Korsgaard, C. (1996a) *The Sources of Normativity* O’Neill, O. (ed.), Cambridge, Cambridge University Press, esp. ch.4; see also Korsgaard, C. (1996b) ‘Kant’s Formula of Humanity’ in her *Creating the Kingdom of Ends*, Cambridge, Cambridge University Press, pp.106-132.

deserve respect provides another explanation of its value – in particular, its ‘symbolic’ value.

The second normative framework in which autonomy plays a central role is in picking out those agents who are candidates for political participation. Those agents who are autonomous are entitled to have their interests and values considered when the political principles, according to which those agents will be governed, are being determined. Insofar as the status of autonomous agent entitles an individual to have their interests and values presented in such processes, we can see why autonomy is valuable. Thus there are constraints on a conception of autonomy: the class of agents identified as autonomous should be coextensive with the class of agents who are entitled to such normative goods.¹⁰ I will not focus in detail on these constraints in the rest of the thesis, leaving these issues for further exploration elsewhere. However, I intend the positive account I outline in the third part of thesis to sit with a conception of autonomy that can feature in these normative frameworks.

1.1.3 Paternalism

The thought, expressed by Griffin, that we would not want another to interfere in our choices even if it were for the sake of our good, is essentially an anti-paternalistic one.

¹⁰ In political philosophy, the processes by which political principles of justice are ratified are sometimes conceived of hypothetically. A conception of autonomy also has a role in identifying and feeding into a theory the features that characterise the participants in the hypothetical process. For instance, on one picture the participating agents are self-interested agents, rationally concerned to secure the pursuit of their interests and values. (See e.g. Gauthier, D. (1986) *Morals by Agreement*, Oxford, Oxford University Press, as discussed by Christman, J. (2005) ‘Autonomy, Self-Knowledge and Liberal Legitimacy’ in Christman, J. and Anderson, J. (eds.) 2005, Oxford, Oxford University Press, pp.330-359.)

On another picture, the agents recognise the rational commitment to respecting the expressed values of others, and hence seek the principles that best enshrine this respect. These differences reflect different conceptions of the traits by which autonomous agents are characterised. Christman attributes this view to John Rawls (see Rawls, J. (1971) *A Theory of Justice*, Cambridge, Harvard University Press; Rawls, J. (1993) *Political Liberalism*, Columbia University Press). Christman discusses these different pictures of the agents that participate in processes of liberal legitimacy, and argues that we should accept the latter; agents lack the kind of self-knowledge and transparency required to take seriously the interests and values from which they bargain on the former view.

This attention to the features of concrete agents is welcome: some theorists have objected that the association of autonomy with independence has led to inaccurate representations of the hypothetically deliberating agents in this process. See Seyla Benhabib (1994) ‘Deliberative Rationality and Models of Democratic Legitimacy’ *Constellations* 1 (1), 26–52. See also Okin, S. M. ((1989) *Justice, Gender and the Family*, Basic books, esp. Chapter 5) on concerns about who is included in Rawls’ deliberative procedures: his original theory included that agents have the property of being heads of households; hence many women in the private sphere were excluded from the decision making process. I do not directly address this issue in the thesis.

Paternalistic interventions are those in which one agent judges that her assessment of what is best for another agent ought to override that agent's own decision or choice. But, building on the thought that autonomous agents deserve 'recognition respect', to override an agent's decision in this way – even with the agent's best interests in mind – is (other than in exceptional cases) a failure to properly respect the autonomy of the agent. As Scanlon claims, "we generally think that the fact that the affected parties chose or assented to an outcome is an important factor in making that outcome legitimate".¹¹ This point has application both with regards legitimacy of certain social and political institutions, and the legitimacy of particular interventions – by the state or individual agents. On occasions in which an agent has not assented to some outcome – even some outcome concerned with her own good – this particular 'legitimising factor' is not present.

Thus (absent strong justification) autonomous agents should be protected from paternalistic intervention. This protection can be understood in a normative sense, and has following form: the burden of justification falls on any proposed intervention with an autonomous agent's decision or action. If no adequate justification can be given, then the intervention is morally criticisable.¹² This normative framework is one that I will focus on in detail in this thesis – the adequacy of relational conceptions of autonomous agency will be assessed in light of their ability to stand in this particular role. The question I will consider in Chapter 6 is whether relational conceptions of autonomous agency can make sense of the claim that certain interventions cannot meet the justificatory burden. As we will see, there are different views on which justifications are adequate.¹³ Can some such interventions be justified whether or not the autonomous agent would assent to them (the hard paternalist view)? Or can only those interventions to which the agent would consent, with further information, be justified (the soft paternalist view)?

Whilst it may be true that "all agree ... that another function of the concept of autonomy is to mark out the parameters within which a person is immune from paternalistic intervention"¹⁴ it remains a live question what kinds of intervention are

¹¹ Scanlon, T.M. (1986) Lecture 1, p.155.

¹² See Gaus, G. (2005) *Op cit.* He argues that a principle that has this form is a core part of liberalism.

¹³ For discussion, see Dworkin, G. (1972) 'Paternalism.' *Monist* 56 pp.64-84; Dworkin, G. (1983) 'Paternalism: Some Second Thoughts', R. Sartorius, (ed.) *Paternalism*, Minneapolis, University of Minnesota Press, pp.105-112.

¹⁴ Christman, J. (2004) *Op Cit*, p.157.

consistent with properly valuing autonomy. I will be primarily concerned with hard paternalism, for the purposes of this thesis. In particular I will be concerned with what I call ‘the core cases’ of hard paternalism; cases which any conception of autonomy should be able to diagnose as unjustified. The central concern here is whether relational conceptions of autonomy can do the work in picking out those agents who should be protected from paternalistic intervention. This is a benefit to which autonomous agents are entitled. Thus, we can see why autonomy, insofar as it creates a normative constraint upon interventions in agents’ choices and actions, is valuable.

1.1.4 Summary

Thus far we have seen that the notion of autonomy plays key roles in a range of normative frameworks: grounding respect, identifying the participants (or features of hypothetical participants) in collective decision making, and – of particular interest in this thesis – protecting from paternalism. The role of autonomy in these normative frameworks means that being categorised as autonomous comes with important normative benefits, and this explains, in part, its value: autonomous agents ought to be respected, ought not have their choices and actions intervened with, and ought to be included in processes that legitimise political power.¹⁵ We can also surmise, from this brief survey of these normative frameworks, that the three roles are intimately connected: that autonomous agents ought to be respected means that their choices and actions ought not (absent strong justification) be interfered with, and that if they are to be subject to political power, this requires justification (which takes the form, in liberalism, of collective rational endorsement).

The problem of how to accommodate, in social and political terms, the extent to which agents should be protected from paternalism, treated with respect, and be ‘free to make up their own minds’ is (part of) what Scanlon refers to as “the political problem of free will”.¹⁶ It is important to distinguish this cluster of problems from that to which Scanlon refers as “the personal problem of free will” which concerns how we might explain the sense in which we ‘own’ some of our actions, rather than feeling “manipulated, trapped, reduced to the status of a puppet” – as we would when we

¹⁵ This is not to say, of course, that autonomy has merely instrumental value in securing these benefits.

¹⁶ Scanlon, T.M. (1986), *Op.Cit*, pp.156, 157.

discover we have been manipulated or hypnotised, say.¹⁷ It is the former set of problems that provides part of the framework for the examination of autonomy in this thesis.

Insofar as being autonomous means being accorded this range of normative benefits it is not difficult to see why it is, quite apart from any judgment about its intrinsic value, that we care about autonomy. We thus get a clearer sense of the structure of the concept: one might be self-governing to a greater or lesser degree. But there will be a relevant threshold above which agents are entitled to the normative benefits set out. The conditions for autonomy are those which must be met for the agent to meet this threshold; she may then manifest the traits required for this to a higher degree (be more autonomous) – but this will not mean that she has more of these benefits (although it may help to make it absolutely clear that she is entitled to them).

In recent decades, the notion of autonomy has received much attention from feminist philosophers.¹⁸ In the following section, I set out what has motivated this attention to autonomy in particular, and the direction that the debate has taken.

1.2 Towards relational autonomy

The attention from feminist theorists has prompted the development of ‘relational’ conceptions of autonomy. A relational conception of autonomy is one which has “amongst its defining conditions requirements concerning the interpersonal or social environment of the agent”.¹⁹ In this section, I clarify the kinds of conditions at issue, spell out some of the motivations for the attention to such conditions, and set in context the project of the thesis.

¹⁷ *Ibid*, p.157. Note that Scanlon is concerned in particular with why ‘the causal thesis’ – namely, the thesis that there are external causes for all of our choices and actions – does not lead us to feel so alienated from all of our choices and actions. I am not concerned with the problems raised by causal determinism here. I address the connections between the debates about relational autonomy and the debates about free will in my paper ‘The Metaphysics of Relational Autonomy’ (forthcoming in *Feminist Metaphysics* Witt, C. (ed.).).

¹⁸ See e.g. Code, L. (1991) “Second Persons” in *What Can She Know? Feminist Theory and the Social Construction of Knowledge*. L. Code. New York, Cornell University Press; Held, V.(1993) *Feminist Morality*, Chicago, University of Chicago Press; see also the essays in *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Self*, 2000, MacKenzie, C. and Stoljar, N. (eds.), Oxford, Oxford University Press.

¹⁹ Christman, J. (2004) *Op Cit*, p. 147.

1.2.1 Clarifying relational conditions

It is important to distinguish the claim that autonomy is relational from another, similar claim, about the relational nature of agents. For example, in their volume, *Relational Autonomy*, Catriona MacKenzie and Natalie Stoljar characterise ‘relational autonomy’ as:

an umbrella term, designating a range of related perspectives... premised on a shared conviction that persons are socially embedded, that agents’ identities are formed within the contexts of social relationships and shaped by a complex of intersecting social determinants such as race, class, gender, and ethnicity.²⁰

However, MacKenzie and Stoljar have here characterised conceptions of agents that are relational, rather than conceptions of autonomy. But there is an important difference between conceiving of agents relationally, and seeing autonomy as relational. One might agree that agents are indeed socially embedded and so on: but these claims may not have anything to do with autonomy. A conception of agents as relational may be accepted, whilst the conditions for autonomy may still fail to reference to the social environment of the agent. Conversely, an individualistic conception of agents might be held in conjunction with a relational conception of autonomy.²¹

So, the question of whether the conditions for autonomy are relational is not the same as the question of whether ‘selves’ or agents are relational.²² The claim about the relational nature of autonomy is my concern here. How, then, might we understand this claim, and what has motivated attention to it?

1.2.2 Motivations for relational autonomy

Some feminist philosophers have expressed discontent with the notion of autonomy, claiming that it presupposes a view of agents that is unrealistic. Lorraine Code, for example, has suggested that the notion privileges a conception of autonomous agents that casts them as essentially independent units, for whom self-sufficiency and

²⁰ MacKenzie, C. and Stoljar, N. (2000) ‘Autonomy Re-figured’ in their *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Self*, Oxford, Oxford University Press, pp.3-34, see p. 4.

²¹ See Friedman, M (2003), *Autonomy, Gender, Politics* New York, Oxford University Press, for a view with this structure.

²² For a discussion of the relational conditions for agents, see e.g. Barclay, L. (2000) ‘Autonomy and the Social Self’ in MacKenzie and Stoljar (2000), pp.52-71. See also Louise Antony, who argues against conceiving selves relationally: Antony, L. (1995) ‘Is Psychological Individualism a piece of Ideology?’ *Hypatia* 10 (3), pp.157-173

detachment from others is the ideal.²³ As characterised by Diana Meyers, this critique holds that “self-governance has been taken to presuppose unfettered independence from other individuals, as well as from the larger society”.²⁴

This has prompted theorists to consider closely the ways in which social relations impact upon autonomy. The thought is that if it turns out that social relations are in some way necessary for autonomy, then it will be clear that the presupposition of ‘unfettered independence’ is false. Now, Marilyn Friedman points out that critiques of autonomy such as Code’s are targeted against something of a straw man. She observes that, in fact, some existing conceptions do take into account the social relations that are required for autonomy. For example, some views draw attention to the fact that “socialization is crucial to the development of the capacity to be a chooser”.²⁵ Likewise, Thomas Hill Jnr. notes that we can value autonomy “without in any way implying that self-sufficiency, independence, and separation from others are goals worth pursuing”.²⁶

There are two points, here: first, no plausible view should deny that there are some causally necessary social conditions for autonomy. As human beings, we are dependent upon others for much of early life for the development of a range of skills –

²³ See Code, L. (1991) *Op.Cit.* Those concerned that autonomy privileges the value of independence point out that many in positions of dependency exercise much careful self-governance. See, e.g. Meyers, D (1987) ‘Autonomy and the Paradox of Feminine Socialization’ *The Journal of Philosophy* 84 (11), pp.619-628.

Other theorists have been attentive to the different kinds of virtues that might be manifested in a purportedly non-autonomous life, arguing for that too much attention to independence and self-sufficiency fails to acknowledge these other important traits. See Gilligan, C. (1994) *In a Different Voice: Psychological Theory and Women’s Development*, Harvard University Press. Gilligan argues that Kohlberg’s model of moral development (See Kohlberg, L. (1984) *The Psychology of Moral Development: The Nature and Validity of Moral Stages*, Harper Collins) fails to recognise the value of the distinctive style of reasoning ‘care reasoning’ that women use. The care perspective is ranked lower on the scale of moral development, whilst ‘men’s’ style of moral reasoning (justice thinking) is the apex of the scale.

For a summary of empirical criticisms of Gilligan’s view, see Saul, J. (2003) *Feminism: Issues and Arguments* Oxford, Oxford University Press. Esp. chapter 7. Saul’s chapter also discusses philosophical problems with Gilligan’s claims. For more on the latter, see Friedman, M. ‘Beyond Caring: The Demoralisation of Gender’ in *An Ethic of Care: Feminist and Interdisciplinary Perspectives*, Laraabee, M.J. (ed.) Routledge, pp.258-275; Card, C. (1995) ‘Gender and Moral Luck’ in *Justice and Care*, Held, V. (ed.) Westview Press, pp.70-100.

²⁴ Meyers, D. (2000) ‘Intersectional Identity and the Authentic Self?: Opposites Attract!’ in *Relational Autonomy* MacKenzie and Stoljar, eds. pp. 151 – 181, see p.152.

²⁵ Friedman, M. (2003), p. 90. She references as views that give some attention to the social conditions that may be causally necessary for autonomy: Benn, S.I. (1982) ‘Individuality, Autonomy and Community’ in *Community as a Social Ideal*, Kamenka, E. (ed.) New York, St Martin’s Press, pp.43-62; Dworkin, G. (1988) *Op.Cit.*; Hill, T. Jnr. (1991) *Autonomy and Self-Respect*, Cambridge, Cambridge University Press.

²⁶ Hill Jnr., T. (1991) ‘The Importance of Autonomy’ in his *Autonomy and Self-Respect*, Cambridge, Cambridge University Press, pp. 43-51, p.49.

including the skills needed for self-governance. If this is all that is meant by the claim that autonomy is relational, then we should accept the view. The second point addresses the compatibility of autonomy and dependency: valuing autonomy “does not deny anyone the choice to share with others, to acknowledge one’s dependency, to accept advice, or even to sacrifice for the interests of others”.²⁷ It seems that the concerns that originally motivated the formulation of ‘relational’ views of autonomy are misplaced: there are causally necessary relational conditions for autonomy, and autonomy is not reserved for only those who are independent and self-sufficient.

But some theorists have argued for a stronger kind of relational view. Paul Benson, for example, argues that causally relational conceptions of autonomy “imply that the character of our interpersonal or social situation ... can only affect our free agency accidentally, by virtue of its potential influences on our capacities.”²⁸ Social relations might impact upon an agent’s autonomy-relevant capacities, but the connection is merely contingent. Indeed, it is important to note that we might accept the claim that there are causally necessary conditions for autonomy, whilst maintaining that the conditions for being autonomous are not relational. That is to say: becoming autonomous might require certain social relations, whilst, subsequently being autonomous does not.

But what might the claim that social relations are constitutive of autonomy amount to? Should we accept the claim, appropriately understood, that some of the constitutive conditions for autonomy make reference to the social environment of the agent? These are the questions that I will be concerned with in this thesis.

1.2.3 The project of the thesis

The question, then, is whether a conception of autonomy should acknowledge not merely causally, but also constitutively necessary relational conditions:

are social relations merely causal conditions that are necessary to bring autonomy about but are external to autonomy proper, rather like sunshine causing plants to grow? Or are they somehow partly constitutive of autonomy?²⁹

²⁷ *Ibid*, p.49.

²⁸ Benson, P. (1994) ‘Free Agency and Self-Worth’ *The Journal of Philosophy* 91(12) pp.650-668, see p. 666.

²⁹ Friedman, M. (2003) *Op. Cit.* p.96.

Marilyn Friedman writes:

This unresolved issue is one major philosophical concern that continues to divide, on the one hand, feminists who advocate a [constitutively] relational account of autonomy from, on the other hand, theorists who acknowledge that social relationships contribute [causally] to autonomy.³⁰

She goes on:

Relational, or constitutively social, accounts of autonomy such as these, I believe, set the stage for the next round of feminist explorations of autonomy. A crucial aim of these explorations should be to determine what it could mean to say that autonomy is intrinsically or constitutively social. Another crucial aim should be to determine whether feminism really needs to regard autonomy as intrinsically or constitutively social.³¹

This thesis undertakes this next round of exploration, critically assessing existing views of what it is for social relations to be constitutive of autonomy (Part I and Part II). Finding these wanting, I set forward, in Part III of the thesis, some constitutively relational conditions for autonomous action. It will be clear that all theorists, not only feminists, need to regard autonomy as constitutively social in this way.

Now, with regards the concern about ‘unfettered individualism’: will adopting a constitutively relational conception make clear the falsity of this presupposition? Even though no theory in fact demands substantive independence, it might nonetheless be that too much independence is a bad thing. So a critique of too much independence might be desirable. Friedman suggests that, whilst this is so, a conception of autonomy will not provide the grounds for such criticism:

a critique of substantively independent behaviour will have to be based on something other than the ideal of autonomy. We cannot fault autonomy theories for failing to do what might lie beyond their proper scope.³²

³⁰ *Ibid*, p.96.

³¹ *Ibid*, p.97.

³² *Ibid*, p.93.

One of the aims of the third part of this thesis is to argue that autonomy does not, when properly conceived, promote or presuppose the value of independence.³³ I shall ultimately argue that not only is dependency compatible with autonomy, it is sometimes required for its exercise. This claim goes beyond the observation, from Hill, that dependency is compatible with autonomy; and, contra Friedman, we will see that the notion of autonomy can provide the basis for a critique of substantive independence as an ideal. This will be argued for in the third part of the thesis.

My project, then, is initially to explore whether there is good reason to accept constitutively relational conditions for autonomy; my initial findings will suggest that there is not. However, in Part III of the thesis, I will argue that there are, in fact, relational conditions for autonomous action. Thus the project, ultimately, might be understood as continuous with that of aiming to ‘refigure’ the notion of autonomy, taking into account the way in which social relations are necessary for autonomy.

1.3 Aspects of autonomy

The concern to take into account the social dimensions of autonomy might focus on different aspects of autonomy. For example, Natalie Stoljar, whose account I examine in chapter 2, looks at the impact of social conditions upon the autonomy of an agent’s particular choices. Paul Benson and Marina Oshana focus on the potential for social relations to thwart autonomous agency (considered in chapter 5). Few theorists explicitly separate out these different dimensions. But the separating of autonomy into its three aspects – agency, choice, and action – is instructive, and central in this thesis, where I consider whether there are relational conditions for each aspect in turn.

Failing to treat each aspect separately may result in some of the conditions for autonomy being missed. We should not expect the conditions for autonomous agency to be the same as those for autonomous choice. The former pertains to the capacities that the agent possesses. The latter pertains to the agent’s exercise of those capacities on a specific occasion. Likewise, the conditions for action will make reference to, for example, the agent’s bodily movements. Such considerations will not be relevant in the same way to autonomous choice or agency. Consider the following examples:

³³ I will, however, accept that it is ‘individualistic’ in a non-problematic way – namely, that it is a property of individuals. See Friedman (2003) *Op.Cit.* (Chapter 1) for discussion of this claim.

Brain Manipulation: suppose that manipulation is an autonomy undermining factor. Gonzalez is an autonomous agent: he has the capacity for autonomous deliberation, choice and action. He is (unbeknownst to him) subject to neural manipulation, such that whenever he seeks to buy a soft drink, he chooses cola (suppose that prior to the manipulation, he would rank other drinks as preferable). With a busy schedule, he has to plan carefully his weekly shop. When he does so, it is clear that he is exercising his capacities for rational deliberation and choice: his autonomy-relevant capacities. When he enacts the plan, and takes from the shelf apples, broccoli, and so on, he chooses autonomously, and acts autonomously. However, when he puts the cola in the basket, his choice of product is not autonomous.

In such a case, we should say that although Gonzalez is an autonomous agent, and is shopping autonomously, his choice for coca-cola is non-autonomous. But having made the choice for this soft drink (non-autonomously), his self-governance in carrying out the choice is exemplary.

Spasm: Suppose that autonomous action, plausibly, requires that one can execute the bodily movements necessary to carry out one's choice. Gina is doing a weekly shop. She too has carefully planned and coordinated when to shop, and what to buy, having exercised her autonomy-relevant capacities. Unlike Gonzalez, Gina has avoided all neural manipulation. When she chooses to buy cola, she does so autonomously. However, when she reaches the soft drinks aisle, she is momentarily struck by a spasm, knocking a four pack of cola into her basket. When she recovers, she continues on to the checkout. Although her choice of product was autonomous, the action of putting the cola in the basket was not autonomous (although the outcome was, incidentally, the same as it would have been had she acted autonomously).

Gonzalez and Gina both experience problems of autonomy, but the problems are quite different, and attach to different aspects of their autonomy. Both are autonomous agents. Each shops autonomously. But one of Gonzalez' particular choices is non-autonomous, although he is otherwise an exemplary self-governing agent. One of Gina's actions is non-autonomous, although none of her choices are. These examples show the way that the three aspects of autonomy come apart. Given the different nature of choice, agency and action and given that, accordingly, the necessary conditions for each aspect will be different, they should be investigated separately.^{34,35} At each juncture, it may be that

³⁴ There is a precedent for this in the literature on moral responsibility, where it is commonplace to consider whether agents are morally responsible, and then consider the further issue of whether, for any particular action or choice, the agent was responsible for that action or choice (the presence of reasonable but false beliefs may mitigate responsibility for a particular choice, say).

³⁵ Note also that this approach is consistent with Hornsby's rejection of those views according to which "characterising self-determined agency is a matter of marking out a special class of events

there are relational conditions for autonomy. But it is important to note that there may be relational conditions for one aspect, and not another.

1.4 A framework

The considerations raised above show us the importance of a conception of autonomy satisfying the following desideratum:

- (A) a conception of autonomy must be able to cohere with the normative frameworks that in part explain its value.

We will also be considering whether we should accept a conception of autonomy according to which:

- (B) the conditions for autonomy are constitutively relational.

Now, if the notion of autonomy is to cohere with the normative frameworks, a further constraint is presented, namely: given intuitions about the class of agents entitled to these benefits, a conception of autonomy must be in accord with these intuitions. Of course, intuitive support is a desideratum for any conception.

It is important to note, though, that there are some specific intuitions that theorists developing relational conceptions have aimed to capture. Some theorists have expressed intuitions to the effect that choosing and acting in accordance with oppressive norms is non-autonomous. Such choices do seem to be problematic in a way that many everyday choices do not.³⁶ Some theorists aim to distinguish between (for example):

[autonomous actions]” (p.10). On this view, autonomous agency is identified by delineating the class of autonomous actions an individual performs. Her concern is that this view fails to make enough room for agents in a conception of self-determined agency. See Hornsby, J. (2004) ‘Agency and Actions’ in Steward, H. and Hyman, J. (eds.) *Agency and Action*, Cambridge, Cambridge University Press, pp.1-23.

Likewise, my strategy is in line with James Stacey Taylor’s rejection of ‘the assumption of transitivity of autonomy’. On this assumption, if a choice is autonomous, the action that follows from it is also autonomous. Taylor argues, with reference to cases of coercion, that this assumption is incorrect. I hope the brief discussion of the examples here also shows the assumption to be false. See Taylor, J.S. (2003) ‘Autonomy, Duress and Coercion’, *Social Philosophy and Policy* 20 (2), pp.127-155, see esp. pp.133-138.

³⁶ See, e.g. Stoljar, N. (2000) ‘Autonomy and the Feminist Intuition’ in *Relational Autonomy: Feminist Perspectives on Autonomy and Agency*, Oxford, Oxford University Press, pp.94-111. See also Benson, P. (1991) ‘Autonomy and Oppressive Socialisation’ *Social Theory and Practice* 17, pp.385-408. It is important to note that Benson’s view in this paper is different to his later view which I consider in Chapter 5.

Troubling: Annie chooses to become a deferential housewife.
Ok: Clyde chooses to become an art dealer.

Those who find Troubling troubling seek to capture these intuitions. Indeed, as we shall see, some have claimed that only a relational conception can make sense of them. The motivation to take into account intuitions is a central aim of some such conceptions, then. So the third part of our framework is:

- (C) A conception of autonomy should cohere with commonsense intuitions about who is autonomous and who is not.

In particular, a constraint is set by the normative frameworks in which autonomy figures: intuitions about who deserves the normative benefits must be respected. At least, any departure from common sense intuition must be motivated and justified.³⁷ These three considerations will be relevant to our assessment of the views I examine, and guide our evaluation of them. With this in mind, I now briefly outline the strategy of the thesis.

1.4.1 Part 1: On relational (substantive) conditions for autonomous choice

One way in which a condition might be constitutively relational is by building substantive value constraints into an account of autonomy. Choices that do not accord with the specified value are not autonomous. Certain social conditions may hinder choice in accordance with the specified value; or certain social options may not be consistent with the value. I consider two explicitly ‘value-laden’ views of this kind: the first from Natalie Stoljar (Chapter 2); the second from Thomas Hill Jr., and defended by Marcia Baron (Chapter 3). Having set out the problems that each of these specific views face, in Chapter 4 I argue that no substantive condition can do the work required of it. Specifically, such a condition is unable to make intuitive discriminations between choices that are intuitively autonomous, and those that are not (see (C), above). This is so despite the stated aim to capture intuitions about agents whose autonomy is thwarted by oppressive social contexts. Thus constitutively relational conditions for autonomous choice such as this remain unmotivated.

³⁷ Because, of course, some feminist theorists have made claims that, at the time, did great violence to common sense intuition. For example, the suggestion that women could work in the public sphere rather than be confined to the home would once have been in tension with common sense intuition. (Thanks to Jimmy Lenman for pointing this out.)

1.4.2 Part 2: On relational conditions for autonomous agency

Constitutively relational conditions might, rather, be understood to demand that the agent stand in certain social conditions or relations. This is the strategy of Paul Benson and Marina Oshana, whose views I consider in Chapter 5. It is here we see the difficulty of offering a relational conception that can stand in the desired normative frameworks; Benson fails to offer constitutively social conditions (so I argue); yet an account such as Oshana's which posits this kind of condition cannot play the normative role of setting the bounds for paternalism. The argument of Chapter 6 is a general one, the conclusion of which is that any constitutively relational conception of autonomous agency will fail in this way; only causally relational conditions for autonomous agency should be accepted.

1.4.3 Part 3: On relational conditions for autonomous action

Now, the kinds of views I consider in the first two parts of the thesis are 'value-laden', in that they incorporate, with the relational conditions, certain values or ideals. This kind of value-laden condition has been identified, by John Christman, with constitutively relational autonomy itself.³⁸ Given the arguments of these earlier chapters, if he is right, this means that no account could both meet both parts of the framework, (A) and (B): that is, no account that is constitutively relational could play the normative roles that account for autonomy's value. But in Part III of the thesis we see that this conflation of relational conditions with value-laden conditions is wrong.

An aspect of autonomy alluded to in many accounts, but rarely explicitly discussed, is that of autonomous action. In this part of the thesis (Chapter 7), I argue that social relations are constitutive of autonomous action. We will see not only that such a view is well placed to identify the social conditions required for autonomous action, but that incorporating such conditions does not also incorporate values or ideals; thus the other normative roles can be fulfilled. I sketch a particular condition that can make good sense, if filled out, of how oppressive social contexts can thwart autonomy. Even in outline form, the work of this chapter should make it clear that a proper understanding of relational autonomy is of importance to everyone, not just, as Friedman suggests, to feminist philosophers.

³⁸ See Christman, J. (2006) Book Review: Marilyn Friedman, *Autonomy Gender, Politics* (2003), in *The Journal of Value Inquiry* 40, pp.137-143.

Chapter 2. Substantive account I: Stoljar

This part of the thesis is concerned with autonomous choice. We can start with a commonsense understanding of autonomous choice as follows:

(CHOICE) an agent's choice is autonomous when the agent exercises self-governance over her deliberations and subsequent choice making.

According to substantive conceptions of autonomy, if an agent fails to choose in accordance with a certain value, she fails to be properly self-governing. Such conditions identify certain social environments as incompatible with autonomous choice: those which are in conflict with the specified value. Before looking at the argument for the substantive relational conception outlined by Natalie Stoljar, it will be useful to set out, in brief, the kinds of accounts to which these views are proposed as alternatives. 'Content-neutral' conceptions of autonomy hold that, with respect to its objects, any choice can in principle be autonomous, so long as the agent has made that choice in a way that meets the necessary conditions on the process or procedures by which she made that choice; or, perhaps on the structure of her motivational or psychological states.¹

In contrast, the kind of relational conceptions with which we are here concerned, namely, substantive conceptions, place constraints not merely on the process, but on the substance of the choice. These conceptions are characterised by Paul Benson as follows:

substantive accounts of autonomy... [are] those in which the contents of the preferences or values that agents can form or act upon autonomously are subject to direct, normative constraints ... there must be some things that autonomous agents cannot prefer or value without sacrificing some

¹ See e.g. Christman, J. (1991a) 'Liberalism and Individual Positive Freedom' *Ethics* 101 (2), pp.343-359; Christman, J. (1991b) 'Autonomy and Personal History' *Canadian Journal of Philosophy* 21, pp.1-24; Dworkin, G. (1988) *Op.Cit.* for procedural accounts. See Frankfurt, H. (1971), 'Freedom of the Will and the Concept of a Person' *The Journal of Philosophy* 68(1), pp.5-20; Watson, G. (1975) 'Free Agency' *The Journal of Philosophy*, 72(8) pp.205-220 for structural accounts. Procedural and structural conditions are not mutually exclusive – Christman's account, for example, posits both procedural and structural requirements.

autonomy, where this restriction depends immediately on the substance of such preferences or values.²

On this view, in addition to the necessary processes, the agent is not properly self-governing unless she also chooses in accordance with certain values.³ The claim is that only an account that incorporates values in this way can make sense of intuitively non-autonomous choices of agents in oppressive social contexts.⁴ Indeed, Natalie Stoljar claims that:

to vindicate the feminist intuition that the subjects [who acted in accordance with ‘false and oppressive norms’] are not autonomous ... feminists need to develop a strong substantive theory of autonomy.⁵

It is quite clear, then, that the aim is to diagnose what is taken to be intuitively problematic about the choices of certain agents, operating in social contexts in which

² Benson, P. (2005b) ‘Feminist Intuitions and the Normative Substance of Autonomy’ in Taylor, J.S. (ed.) *Personal Autonomy*, Cambridge, Cambridge University Press, pp.124-142, p.133. The kind of substantive account that is characterised here is often referred to as a strong substantive account. Others, which do not place constraints directly on the choice, but rather build normative content into an account of autonomy elsewhere, are called ‘weak substantive’ accounts. An instance of this account is considered in Part II, although I avoid this terminology throughout the thesis.

Note also that Benson talks of preferences or values. Thus it is not clear whether we should understand this characterisation as pertaining to non-autonomous choice (choice informed by those preferences or values) or non-autonomous agents (who hold those preferences or values). I will understand his claim as pertaining to choice, although my claims should hold concerning agents also. As we will see, the kinds of claims from Stoljar and Hill Jnr. make this focus on choice natural. Whilst there is rarely clarity about which aspect of autonomy is under consideration, I think this is how best to interpret their claims.

³ Those who offer substantive constraints often say little about the precise nature of the content-neutral conditions to which they envisage the substantive condition be added. Stoljar, for example, runs through various procedural and structural conditions, showing none of them to capture intuitions about the autonomy of the individuals in her examples. But she does not say whether she endorses any of them as part of a set of jointly sufficient conditions.

⁴ A narrower understanding of substantive conditions can be found in the literature. Marina Oshana (2006, *Personal Autonomy in Society*, Hampshire, Ashgate Publishing Ltd, p.41) endorses the understanding set out (but not endorsed) by Marilyn Friedman: an agent is autonomous “only if she choose[s] in accord with the value of autonomy itself, or, at least, choose[s] so as not to undermine that value” (2003, *Op.Cit* p.19). I set aside this understanding here for the following reasons: first, the arguments I consider here incorporate values other than that of autonomy. Second, the argument I give in chapter 4 addresses conditions that demand choice in accordance with any value – this will include the value of autonomy as well as the others considered – so a broader understanding of substantive condition both represents the literature, and serves our present purposes, better. Third, given that the theorists say little about which content-neutral conditions the substantive condition is intended to supplement (see footnote above), it is not clear exactly what is meant by the claim that agents must choose in accordance with the value of autonomy. If it means that the agent must choose in accordance with (the value of) the content-neutral conditions, this will not provide the kinds of restrictions that substantive theorists seek. Choices for oppressive modes of behaviour, or for roles of deference, substantive theorists concede, can meet all the content-neutral conditions.

⁵ Stoljar, N. (2000) *Op.Cit*.p.109.

gender inequality exists, and in which gender socialisation is deemed to have oppressive effects.⁶

In the following I set out an example of the choices that trigger Stoljar's 'feminist intuition', and outline the content-neutral conditions that she takes to be inadequate in diagnosing the (purported) non-autonomy of the choices. Stoljar merely sketches the kind of relational condition she envisages; I will argue that there is little reason to suppose that a more detailed filling out of this sketch would do the work required of it.

2.1 Stoljar's 'feminist intuition'

Stoljar draws on a set of interviews with women who had taken contraceptive risks, and subsequently sought abortions.⁷ In these interviews, the women reported on the reasons for which they had engaged in such risky behaviour, and the author of the study, Kristin Luker, argues that, having considered the reasons offered, we should conclude that these women were not simply foolish or irrational. Rather, she claims, they had "engaged in a process of tacit bargaining with themselves over the costs and benefits of using contraception",⁸ and the choice to take a risk won out. Some of the costs involved were those of informal sanctions that might attach to behaviour that did not conform to a set of social norms concerning sexual behaviour. These norms, as reported, include:

- (n1) it is inappropriate for women to have active pre-marital sex lives
- (n2) women should not plan for or initiate sex
- (n3) fertility increases a woman's worth
- (n4) women are only valuable as wives if they can bear children⁹

Accordingly, insofar as these norms and the likely costs of norm violation were taken into account in deliberations, these women should be understood as rational bargainers rather than irrational risk-takers. An example (based on the study that Stoljar discusses) will be useful for getting clear on the kinds of deliberation and choice that are at issue here (I note in the example where some of the norms are playing a role in her deliberations):

⁶ The claims about oppressive socialisation and oppressive social contexts need not be restricted to the dimension of gender.

⁷ See Luker, K. (1975) *Taking Chances: abortion and the decision to contracept*, Berkeley, University of California Press.

⁸ Stoljar, N. (2000) *Op.Cit.*, p.96.

⁹ *Ibid.* p.99.

Deliberation: Suppose that Almaz, who has had an upbringing in a strongly Catholic community, has been seeing her partner Brad for a few months, and is considering the possibility that they might have sex sometime soon. She wonders whether to seek out some contraception. The following thoughts occur to her in her deliberations:

I could go to the doctors and ask for a prescription for the contraceptive pill. If I go to the doctors and ask for a prescription, the doctor will think I'm having pre-marital sex. And this is a small community; others might find out too... there'd be a lot of talk... [n1, n2]. I'd feel a bit embarrassed and perhaps ashamed to ask, though that shouldn't stop me from asking.

Anyway, the Church says that pre-marital sex is wrong [n1]; and that contraceptive use is wrong; most people around here think the same. I don't want to face censure from them.

But the risk... the worst case scenario would be getting pregnant. But if that happened we could get married perhaps. Settling down and having children would be one way of sorting things so it'd all turn out ok. And it's not certain I'd get pregnant anyway.¹⁰

In this example, we see that the norms about premarital sex and women's sexual agency are operative in her community, such that violating these norms – by having pre-marital sex, and using contraception – would be met with social censure. These norms are supported by religious authority, as well as being widely socially accepted. Thus concerns about the informal sanctions or censures that she might face, were she to seek out contraception, figure prominently in Almaz' reasoning about what choice to make. Seeking out contraception would be seen to violate the norms that many in her community adhere to. The costs of getting contraception, when weighed in the balance, are given weight such that not getting contraception – taking a risk – appears to be the less costly option, and this option is chosen. (For simplicities' sake, I have focused on norms n1 and n2 in the example above, but we can imagine how norms n3 and n4 might serve to make the choice she makes – of taking a risk – more palatable: becoming pregnant will demonstrate her fertility, and hence worth as a woman; her value as a wife will be proven.)

Now, it is clear that Almaz' reasoning is not an exemplar of practical reasoning. We might be strongly inclined to say that she, and the women like her in the study, made a mistake in the weighting of the outcomes, say. But given the beliefs held about the outcomes and the values placed on these, it is not implausible to hold that Almaz is rational – at least to the degree that many of our choices are: that is, imperfectly,

¹⁰ This example is re-constructed from the examples that Stoljar discusses.

rational. Indeed, Stoljar acknowledges that the choices that these women made might be rational. However, she claims that there is a ‘feminist intuition’ that the women in the study, like Almaz, are nonetheless “non-autonomous because they are overly influenced in their decisions ... by stereotypical and incorrect norms of femininity and sexual agency”¹¹ – norms such as (n1-n4).

Now, the norms set out above (n1-n4), are clearly criticisable from a feminist point of view, as Stoljar claims. But that they are operative in the women’s social context means that it is not irrational for the women who will be judged by them to consider such norms, and the costs of non-compliance, in their decisions about contraceptive use. What it does mean, Stoljar holds, is that the women who are influenced by these norms in their choices to engage in risky contraceptive behaviour are non-autonomous in so choosing.

There are two steps in Stoljar’s argument: first, that women like Almaz are non-autonomous in their choices; second, to explain the non-autonomy of the women in their choice-making, a strong substantive account is required. One might reject the first step. Indeed, Paul Benson reports having quite different intuitions about the women in the study and their choices. Moreover, he objects to the thought that there is such a thing as ‘the’ feminist intuition, suggesting that a diverse range of views on the matter would be consistent with feminist commitments.¹² For present purposes, I grant Stoljar the claim that the women are non-autonomous. I want to consider whether even if we grant this, her condition can succeed where, she claims, content-neutral accounts cannot. First, I set out, briefly, the kinds of conditions in the existing literature that Stoljar argues cannot make sense of the non-autonomy of choices the women whose choices were informed by oppressive norms, like Almaz’. This serves as an overview of some of the prominent views in the recent literature on autonomy, and we see in more detail what motivates the relational view propounded. However, I will suggest that content-neutral conceptions may yet have the resources to make sense of her intuition.

2.2 Some content neutral conditions

The following content-neutral conditions may be necessary for autonomous choice. But, Stoljar argues, they cannot be sufficient conditions; the agents in the study meet these

¹¹ Stoljar, N. (2000) *Op.Cit.* p. 98.

¹² See Benson, P. (2005b) *Op.Cit.* The intuitions that Benson has are consonant with *his* conception of autonomy, which I consider in chapter 5.

conditions but are nonetheless non-autonomous. In setting out her arguments, I suggest that we should not be so quick to reject some of the content-neutral conditions as unable to diagnose as non-autonomous the women's choices. I then turn to her proposed substantive condition.

2.2.1 Structural:

Structural conditions for autonomy demand that the agent's choice stems from a motivational set that has a certain structure. One kind of structural view posits conditions that pertain to the coherence of the agent's other motivational states with her choice. A second view focuses on the higher order attitude that an agent takes towards her choice.

a) Coherence.

On this view, an agent's autonomy is a function of the coherence of a particular choice with the rest of her motivational and mental states. If a choice is 'well-integrated' into the agent's motivational set, then she is autonomous in so choosing. What it means to be 'well-integrated' varies according to different accounts. One version of this condition is Christman's, who demands that there be "no manifest inconsistencies among the beliefs and desires in the set of beliefs and desires that contribute to the processes of reflection".¹³

Stoljar claims that the women in the study would, for the most part, meet Christman's condition: the subjects in the contraceptive study do not show inconsistency to the extent that it would count as a "manifest violation of the internal coherence condition".¹⁴ That is to say, there is no case in which the agent's obviously believe, or desire, p and $\neg p$. Stoljar's claim is that the agents do not have conflict which "constitutes a manifest breakdown in the capacity for critical reflection" and this seems right – the agents are not conflicted to the extent that they lose their capacities for critical reflection.¹⁵

¹³ Christman, J. (1991b) *Op.Cit.* See Stoljar's discussion in her (2000), *Op.Cit.* p.103. A stronger coherence condition can be found in Arpaly, N. and Schroeder, T. ((1999) 'Praise, Blame and the Whole Self' *Philosophical Studies*, 93 (2), pp.161-188), who demand that we look not only to those beliefs and desires that are utilised in the process of reflection, but to *all* of the beliefs and desires in the agent's motivational set – not just those that enter into the deliberative process.

¹⁴ Stoljar, N. (2000) *Op.Cit.* p.104.

¹⁵ Stoljar, N. (2000) *Op.Cit.* p.105.

However, it does appear that the agents are somewhat ambivalent. But rather than demonstrating faulty agency, this appears to demonstrate acknowledgement of the tension that arises from the norms. Consider Almaz's reasoning again: she considers the pros and cons of taking a risk, or taking precautionary steps. Some of her beliefs and value commitments pull in conflicting directions: her beliefs about the authority of the church, say, and the views of those in positions of social authority (such as the doctor, who she might feel "ashamed to ask") are in tension with her self-interests – she acknowledges that feelings of shame "shouldn't stop [her] from asking", and that "the worst case scenario would be getting pregnant".

The demands of the norms appear to be in tension with Almaz' self interest, and hence she is ambivalent about which to give precedence to. This seems reasonable, given the set of norms that she has to take into account; it seems that some ambivalence is not always problematic, and that such ambivalence or conflict in the face of conflicting norms is understandable and perhaps justified.¹⁶

b) Hierarchical,

A hierarchical account focuses on the distinctive capacity of persons to take a reflective stance towards the motives they find themselves with. The hierarchical view appeals to the role that an agent's higher order endorsements play, in resolving conflict of lower order desires. When an agent finds herself with a manifest conflict of desires at the first order, say, she can nonetheless be autonomous if she chooses to act in accordance with one of these conflicting desires. Her autonomy depends upon the conflict being resolved at a higher order: insofar as at (say) the second order, one (but not both) of the conflicting (first order) desires is the object of a second order desire, then this is sufficient for autonomy in choice and subsequent action. For example, if an agent desires to smoke and desires not to smoke, and upon reflection finds herself endorsing the desire to smoke, then when she chooses to smoke, she is autonomous in her choice. This is so despite the conflict of first-order desires, and even if her addiction is such that

¹⁶ See Westlund, A. ((2003) 'Selflessness and Responsibility for Self', *The Journal of Philosophy* 112 (4) pp.483-523) for a similar claim to the effect that ambivalence in deferential individuals, with respect to their deference, seems to temper judgments about their non-autonomy. Such claims challenge the thought, expressed by Frankfurt, H. ((1999) 'The Faintest Passion' in his *Necessity Volition and Love* Cambridge, Cambridge University Press, pp.95-107) that ambivalence always poses a problem for autonomy. Marina Oshana, also, expresses doubts about when ambivalence is a threat to agency. Ambivalence, she holds, is only a problem in extreme cases: "conflicted values, desires, and kindred psychological states curtail autonomy where ... they engender conflict of a sufficiently comprehensive kind to render any semblance of an agent impossible." (2006, *Op.Cit*, p. 31). This kind of conflict, we have seen, is not present in Almaz or the women like her in the study.

she could not choose otherwise. The agent's higher order endorsement resolves the conflict in a way that authorises her subsequent choice and action as one that is, in a significant sense, her own.¹⁷ Such a view has been influentially argued for by Harry Frankfurt.¹⁸ Likewise, if an agent chooses to take contraceptive risk, then insofar as she reflectively endorses this choice, she is autonomous in so choosing, despite her first order ambivalence.

According to Stoljar, the women who have been influenced by oppressive norms show that a motivational structure of higher order endorsement that resolves lower order conflicts cannot be sufficient for autonomous choice. Some of the agents meet the hierarchical endorsement condition, she claims, but nonetheless the intuition remains that they lack autonomy in choice. Stoljar claims that the women's choices to take contraceptive risk are, in many instances, endorsed at a higher order desire and so they meet the Frankfurtian endorsement condition. Consider the following first order desires:¹⁹

(d1) desire [not to plan for sexual activity]

Insofar as seeking contraception is a kind of planning for sexual activity, this implies:

(d2) desire [not to seek contraception]

But the agent also desires:

(d3) desire [not to take contraceptive risk]

Given the agent's belief that she is not in possession of contraception, and that she will not find contraception by other means than seeking it out herself, there is a conflict in first order desires here. Her desire not to take contraceptive risk and her desire not to seek contraception cannot both be satisfied. It is this kind of conflict of first order desires that the higher order endorsement can resolve – by 'authorising' one of the conflicting desires.

¹⁷ The notion of authority will come to the fore in part III of the thesis. Note here the emphasis is on intrapersonal authority, to which I return in more detail in chapter 5. with an examination of Paul Benson's view. Later I consider interpersonal authorising and its role in autonomous action.

¹⁸ Frankfurt, H. (1971) *Op.Cit.* Note that Frankfurt makes clear that he intends his conditions to be sufficient but not necessary for autonomy (see p. 19).

¹⁹ It is important to note that the sense of desire at work in Frankfurtian accounts is that of *dispositional* desire, according to which one has a desire if one is motivated to act. Thus one may, in performing an action A, desire to do A in the dispositional sense, even if one has no phenomenal desire (urge, say), to do A. Noting this may help make plausible the claim that Almaz desires not to seek contraception – this does not require that she has a phenomenal urge.

Now, Stoljar thinks it plausible to understand the women as having in their motivational sets higher order endorsements of both of these desires. Suppose that Almaz is certain she does not want to get pregnant; it is plausible then to hold that she endorses the desire not to take contraceptive risk. But she really does give weight to the costs of seeking contraception, such as being held in ill-repute by members of her community whom she esteems. So she endorses the desire not to seek contraception, and thus has the following higher order desires:

- (d2*) desire [desire [not to seek contraception]]
- (d3*) desire [desire [not to take contraceptive risk]]

Insofar as there is endorsement of the first order desires, Stoljar claims, there is no conflict between the first order and higher order desires. Their motivational sets contain endorsements such that “both first order desires are endorsed at a second-order level”.²⁰ According to Stoljar, if Almaz were to act upon her desire (d2), as endorsed by (d2*), she would meet the sufficient conditions for autonomy, as set out by the hierarchical account. She maintains that because there is higher order endorsement, she and women like her in the study, can be regarded – by the lights of the Frankfurtian condition – as autonomous. And this diagnosis, she claims, is wrong.

c) Problems

We should not accept Stoljar’s rejection of the hierarchical condition as sufficient for autonomy, for Frankfurt’s view in fact can make sense of the non-autonomy of the women’s choices. We have seen that it is plausible to hold that the women experience conflict between first order desires. This conflict at the first order of desires requires resolution at the second order. In endorsing both of these first order desires at the second order (with (d2*) and (d3*)), there is no resolution of the conflict; it has merely been elevated a level. There may be endorsement of the first order desires at a higher order, but the higher order desires of the agent (d2*) and (d3*) themselves conflict. We cannot say that (d2) is endorsed in the relevant way, because there is both a desire in favour of it – (d2*) – and a desire that repudiates it, (d3*).

Now, Stoljar notes that this reiteration of conflict at a higher order poses problems for Frankfurt’s account – the well-documented regress of levels of endorsement looms. However, she must also acknowledge that the agents do not meet

²⁰ Stoljar, N. (2000) *Op.Cit.* p.106.

Frankfurt's sufficient conditions for autonomous choice. Stoljar has not done enough to show that the hierarchical structural conditions for autonomy cannot account for the agents who have been 'overly influenced' by the oppressive norms. If such a view can account for the intuition – and I have argued that it can – then part of the motivation to adopt the substantive relational condition is undermined.

2.2.2 Historical

Another well-documented problem with this kind of hierarchical structural account however, is that it fails to take into account the histories of the agent's motivational set. An agent who meets the structural conditions could do so as a result of having this structure induced by a clever neurologist or hypnotist or whatever intuitively autonomy-undermining entity is posited.²¹ This is counterintuitive, and prompts the addition of historical conditions. The historical condition posited by Christman, demands that:²²

- i) the agent was in a position to reflect upon the processes involved in the development of the desire.
- ii) the agent did not resist the development of the desire when attending to this process of development, or would not have resisted, had he attended to that process.
- iii) the lack of resistance to the development is not due to the influence of factors that inhibit self-reflection²³

This kind of condition is intended to deal with the problems that face hierarchical accounts: agents whose motivational structures have been induced by hypnosis and the like will, most likely, resist the formation of a desire brought about by this process of development if they attend to it – or would have resisted the formation of the desire, had they attended to it. If an agent resisted, or would have resisted, the formation of a desire in this way, then this accounts for their lack of autonomy when choosing on the basis of that desire at a later time.²⁴ Can such an account speak to Stoljar's intuition about the non-autonomous choices of the women in the study? Stoljar holds that the agents would not have, even had they reflected upon this process, resisted the

²¹ See Slote, M. ((1980) 'Understanding Free Will' *The Journal of Philosophy*, 77 (3) pp.136-151), for discussion of such problem cases.

²² See Christman, J. (1991a) and (1991b), *Op.Cit.* for argument for and defence of these conditions.

²³ See Christman, J. (1991a), p. 347.

²⁴ An agent may sometimes endorse preferences formed due to processes such as hypnotism: for the sake of forming preferences not to smoke, for example. Insofar as, upon reflection, they do not resist this, then they can be autonomous with respect to this preference.

development of the desires or values that now inform their choices. Her claim is difficult to make sense of, so I set it out here. She claims that:

it is unlikely that they would have [resisted the formation of these preferences when reflecting upon the process of development] precisely because the habits of deference and the internalised norms, that is, the values that govern the agent's motivational structure, would themselves justify holding the relevant desire.²⁵

Stoljar's claim seems to be this: suppose that Almaz has internalised habits of deference (or values that encourage this) and a set of norms (n1-n4). This means that, when she is undergoing the process of developing relevant preferences (the preference not to seek contraception that informs the choice for contraceptive risk, in this case) she would not resist the process. This is because the values and norms (n1-n4) are internalised, and govern her preference formation and subsequent choice. The habit of deference – presumably, a habit to avoid being too critical or free-thinking in making choices – consolidates the role that the internalised norms play in the formation of the preference. So had Almaz thought more carefully about the process by which she came to her preference and choice, she would not have resisted it.

Stoljar contends that the agent might be “in the grip of feminine norms about the dependency of women on men in sex, pregnancy and marriage”²⁶ such that she does not, and would not have, resisted the formation of the preference and subsequent choice to take contraceptive risk. Indeed, Stoljar writes that “the fact that the norms are internalised blocks the capacity of the agent to resist the development of preferences based on the norms”.²⁷ So the agent does not resist, and would not have resisted, the formation of the preference. So it appears that she meets Christman's historical condition.

a) Problems

It is not clear, however, that Christman's historical condition is unable to account for Stoljar's intuition that the resulting choices are non-autonomous. The reason for which the agents would not have resisted, Stoljar claims, is because they were ‘in the grip’ of the norms that governed their deliberations. Might being ‘gripped’ by norms in this way mean that the agent in fact fails to meet Christman's conditions? Christman's historical

²⁵ Stoljar, N. (2000) *Op.Cit*, p.101.

²⁶ *Ibid*, p.102.

²⁷ *Ibid*, p.101.

condition, as he states it, requires that the agent “was in a position to reflect upon the processes involved in” the development of the preference.²⁸ According to Christman, this kind of reflection involves:

being in a position to focus on the processes and conditions that led to the adoption of that desire. That is, a relevantly full description of the steps of reasoning or causal processes that led her to have this desire is available for her possible consideration. This reflectiveness assumes that the agent can become aware of the beliefs and desires that move her ... [that is] to bring to conscious awareness a belief or desire, and concentrate on its meaning.²⁹

(By this latter claim, we should take Christman to mean something more like ‘significance’, rather than just the literal meaning of the propositional object.) If the agent is in the grip of a norm, then it may be that the agent is not in a position to adequately reflect upon the processes by which she forms the relevant desire or preference – perhaps the grip of the norm is such that she is unable to bring it to consciousness and concentrate on its meaning. As characterised by Stoljar, then, it is not clear that the agents would meet the historical condition that Christman sets out.

Moreover, insofar as the internalisation of the norms has blocked the agent’s capacity to resist, then it appears that she will not meet Christman’s third condition, which pertains to the explanation for the lack of resistance. Any lack of resistance to the formation of the preference, according to Stoljar’s description, is not due to the endorsement of the preference; rather, it is due to the ‘blocked capacity’. So described, Christman’s historical condition seems to be able to diagnose the agents’ choices as non-autonomous. The agents ‘in the grip of norms’, or for whom the preference formation is ‘irresistible’ fail to meet Christman’s historical condition.

However, it seems to me that Stoljar need not have made the strong claims about the women such as Almaz being ‘gripped’ by the norms that informed their choices. Rather, on a more straightforward understanding of the women’s deliberation and choice, it is plausible that Christman’s conditions will be met by some of the women in the contraceptive studies. What Stoljar should have said with respect to these cases is the following: because the process of preference and value formation at issue is run of the mill socialisation, and deliberation on the basis of the deliverances of this, it

²⁸ Christman, J. (1991a) *Op. Cit* p.347.

²⁹ *Ibid*, p.347.

is unlikely that the agents would have resisted this process of preference formation. There is no reason to suppose that resistance is warranted with respect to these processes: most of our preferences are formed in this way, and there is little intuitively problematic about this.³⁰

Whilst we should reject Stoljar's explanation of why Christman's condition will be met by the women in the study, then, we should nonetheless accept her conclusion that his account will most likely fail to diagnose at least some of the women in the study as non-autonomous.

2.2.3 Competence: reflective inhibitors, self-knowledge

One of the concerns that Stoljar seems to be getting at with her worries about the historical view is that the conditions under which the agents form preferences or appeal to them in deliberation may be such that the agents are insufficiently competent to properly reflect upon them. Agents may have developed preferences in conditions that inhibit reflective competence, for example. Two versions of such competence conditions are discussed by Stoljar.

a) Reflective inhibitors

Preferences might be formed in conditions that prevent the agent from properly reflecting upon the preferences or values that she is forming. For example, a woman who has formed preferences for the role of housewife in a context in which any other role is regarded with disdain and faces sanction may fail to consider or form any preferences for other options. This failure to consider other preferences inhibits reflection upon this preference – because no other 'live' option was on the table for comparison, say. Subsequent choices in which such a preference figures will inherit this problematic lack of reflection. Such is an example of external factors inhibiting the agent's reflective competence.

³⁰ Of course, socialisation itself might be thought to be problematic. In the literature on free will, there is much literature addressing the fact that we are not – due to such necessary socialisation – responsible for our character – the issue of our responsibility for the choices and actions that flow from it then arises. See Wolf, S. ((1980) 'Asymmetrical Freedom' *The Journal of Philosophy*, 77 (3), pp.151-166) for a good overview of the problem, and her proposed way of dealing with it. I do not address this issue here, but rather start from the thought that most of our everyday choices are autonomous in a way that grounds respect, anti-paternalistic principles, and so on.

Internal factors, such as addiction, may similarly impair an agent's reflective abilities. They may prevent an agent from exercising reflective competence in evaluating her desires, as well as interfering with the role they might play in deliberation. In these cases, the choices that the agent makes on the basis of those preferences should not be deemed autonomous. This kind of 'reflective competence' condition is incorporated into Christman's account.

Now, on the one hand we have seen Stoljar's comment that the women may be 'in the grip' of norms in a way that inhibits reflection on the preference formation. In assessing this competence condition, however, Stoljar claims that we should not think of the women in Luker's study as suffering from defects in reflective competence in such a way as to impair their autonomy. Her claims here seem to be right; and indeed, I have suggested that Stoljar need not make strong claims about the agent being so 'gripped' by the norms. Rather, the agents demonstrate a level of reflective competence. Recall Almaz and the process of deliberation that we supposed she undertook. This process involved reflecting on the options, and their consequences:

If I go to the doctors and ask for a prescription, the doctor will think I'm having pre-marital sex. And this is a small community too, and others might find out too... there'd be a lot of talk... I'd feel a bit ashamed asking, though that shouldn't stop me from asking.

It certainly does not appear that Almaz is prevented, by external factors, from considering options other than taking a contraceptive risk. Indeed, in deliberation she considers the pros and cons of doing so in comparison with the alternatives. Neither is she impaired by internal factors; we need not think that she is gripped by social norms any more than the rest of us usually are.

As Stoljar suggests, whilst the conditions in which preferences have been formed are far from ideal, it is implausible to suppose that the women's critical capacities are seriously impaired. Indeed, that we should consider the women – in line with Luker's recommendation – as rational bargainers, albeit imperfectly rational, requires that we acknowledge the reflective capacities of the women. This conclusion seems right.

b) Self-knowledge

There might be other ways in which an agent's competences as a deliberator can be impaired, such that although she remains rational, she is nonetheless non-autonomous. One such way might pertain to the agent's self-knowledge. Such a self-knowledge condition for autonomy is endorsed by Diana Meyers.³¹ Stoljar suggests, for example, that an agent should not be self-deceived about her preferences; nor should she be self-deceived about the process by which her preferences developed.³² The women in the study, for example, may be self-deceived in failing to acknowledge themselves as sexual agents. Perhaps in not wishing to seek contraceptive, Almaz is failing to acknowledge certain desires that she has (for sex) and therefore lacks self-knowledge.³³ If lack of self-knowledge undermines autonomy, then this lack will explain the intuition that these women's choices are non-autonomous.

Note that there are two ways to understand the demand that agents meet the self-knowledge condition with respect to their preferences (neither Stoljar nor Meyers make clear this distinction). First, it might demand that the agent know what her preferences are. If an agent has preferences that she does not believe herself to have, then she lacks the self-knowledge required for autonomy. This lack means that, for instance, she cannot make choices that take into account how best to satisfy such preferences. Second, a weaker version of the self-knowledge condition might require that an agent has no false beliefs about her preferences.³⁴ This is consistent with her failing to believe that she has certain preferences that she in fact has. For example, an agent might fail to believe that she has a preference for sexual relations (when she in fact does). But this does not mean that she believes that she has a preference for not having sexual relations. On this second way of understanding the self-knowledge condition, the women may fail to recognise that they have certain preferences and values, but still meet the condition.

³¹ See Meyers, D. ((2002) *Gender in the Mirror: Cultural Imagery and Women's Agency* Oxford, Oxford University Press. Esp. Chapter 1), for a competence account that places emphasis on the agent's self-knowledge and abilities to introspect.

³² See Christman, J. (1991b) *Op.Cit.* and Stoljar, N. (2000) *Op.Cit.* at 102 for discussion.

³³ See Anderson, E. ((2002) 'Should Feminist's Reject Rational Choice Theory?' in *A Mind of One's Own* Witt, C. and Antony, L. (eds.) Westview Press, pp.369-387) for this claim.

³⁴ Of course, this issue could be complicated by considering the case of an agent who lacks knowledge of preferences that she does have, and also believes 'that I know about all of the preferences I have'. This entails that she believes 'that I do not have preference for p, for q' and so on. The distinction between the case of failing to have true beliefs, and having false beliefs, might then collapse. However, most agents are aware that they may lack full knowledge of their preferences, so I set aside this wrinkle.

The first way of understanding the condition is too demanding: it is surely the case that we have many preferences that are not readily available for epistemic access, and that we accordingly fail to attend to in deliberation and choice. For example, a short while ago I discovered that a fruit called guanabana exists. Now I presume I might like to try it, although I do not know if I would like it. Until I try it I shall not know. But surely this lack of self-knowledge does not mean that I lack autonomy in choice until I ascertain what my preferences are – even in fruit related choices. If the self-knowledge condition were to demand that agents knew all of their preferences, then many agents like myself and like Almaz will fail to meet the necessary condition for autonomous choice, and the kinds of distinctions that Stoljar is aiming for could not be drawn. The second, ‘no false beliefs’, condition, then, is the more plausible one.

This second way of understanding the self-knowledge condition is not discussed by Stoljar. She argues that the women in Luker’s study may be deceiving others about the extent to which they violate norms that govern their sexual activity, but, she claims, they do not deceive themselves. If she is right about this, then the self-knowledge condition cannot deliver the verdict about the women’s autonomy that Stoljar believes is required. It seems that Stoljar is right: in contemplating getting contraception, Almaz is acknowledging her desire for sex. However, a stronger case is available to Stoljar. Even if the women fail to acknowledge to themselves their preferences for sexual agency they may nonetheless meet the condition on its most plausible reading. If the women fail to believe that they have preferences for sexual agency they might still meet the self-knowledge condition insofar as it demands only that the agent have no false beliefs about her preferences (rather than that the agent have, for each of her preferences, true beliefs about it).

Indeed, Stoljar notes that some – although not all – subjects in the study do acknowledge their preferences for sexual agency, although they also acknowledge that they should not be seen to be sexual agents. For instance, one woman reported: “[I didn’t get a prescription] Because of my father ... we live in a small town ... and I couldn’t go to another doctor without his finding out”. Of this woman, Stoljar writes: “although the agent is influenced by norms of sexual agency to the extent that she wishes to be seen to be observing them, she is not deceiving herself”.³⁵ Thus at least some of the agents about whom Stoljar has the feminist intuition (namely, that they are non-autonomous) meet this weaker self-knowledge condition. Absent good reason to

³⁵ Stoljar, N. (2000) *Op.Cit.*p.103.

accept the stronger version of the condition, it is plausible that the self-knowledge condition could be met by the women. This kind of condition cannot account, then, for the intuitive non-autonomy of the agents.³⁶

2.2.4 Summary

One aim of this discussion was to survey some of the content-neutral conditions that are offered, and to which the relational conditions are supplemented. These kinds of conditions – structural, historical, competence – are candidates, according to the content-neutral theorist, for what brings an agent to the threshold for autonomy; beyond which the agent is entitled to certain normative benefits; for example, the entitlement to have her choices protected from paternalistic intervention (absent strong justification). Stoljar's claim is that none of these conditions can adequately make sense of the intuitive non-autonomy of the women who chose in accordance with oppressive norms. I have suggested that whilst most of these conditions appear to be met by the women – the coherence, historical, and both competence conditions – it is not clear that the hierarchical condition is met. Thus content-neutral conditions may yet be able to account for Stoljar's 'feminist intuition'. If this is so, then much of the motivation for adopting a substantive, relational condition is removed. I now turn to consider the proposed substantive condition. If this condition is plausible, then there may be reason to accept it independently of the adequacy (or not) of the content-neutral conditions. However, I will argue that it is not.

2.3 Stoljar's substantive condition

Recall that Stoljar's aim is to develop a conception that takes into account the social aspects of agency, and explains in particular 'the feminist intuition', according to which those choices influenced by oppressive norms are non-autonomous. Having argued that no content-neutral conception of autonomy can achieve this, Stoljar sets out the structure of the substantive conception she thinks is required. Stoljar claims that it is simply because the agents "are motivated by oppressive and misguided norms that are internalised as a result of feminine socialisation" that they are non-autonomous.³⁷ She endorses the claim that:

³⁶ Note that in Chapter 4, I raise doubts about whether we should accept that there are epistemic conditions for autonomy. In light of this, the conclusion of this section might be read as: if there are any self-knowledge conditions, then at least some of these women meet them.

³⁷ Stoljar, N. (2000) *Op. Cit.*p.98.

the effective internalisation of false or irrelevant norms together with the fact that such norms are false or irrelevant diminish or extinguish agents' capacities for autonomy with respect to decisions governed by the norms.³⁸

The problem for an individual who has internalised and chooses on the basis of such a norm, she claims, is that their deliberation and choice “is informed by oppressive and misguided norms and hence is not autonomous”.³⁹ We can make sense of this non-autonomy, she claims, by placing substantive constraints on the contents of autonomous choice. This builds in a relational dimension to the conditions: insofar as the object of choice is for an option, practice or relationship that is inconsistent with the substantively specified value, the choice is non-autonomous.⁴⁰ Can such a condition deliver the verdict she seeks about the lack of autonomy of such women?

2.3.1. A constraint: false and oppressive norms

As Stoljar admits, her proposal is underdeveloped. Indeed, as it stands, there are two ways of understanding her claims. One reading is suggested by the comment that “because of the internalisation of the [false and oppressive] norm, they do not have the capacity to perceive it as false”.⁴¹ On this reading, her account is a kind of competence condition, which, in fact, could be content neutral. On such an account, anything could, in principle, be chosen autonomously insofar as it is possible for the agent to adequately reflect upon the choice. Some of Stoljar's claims suggest that the problem arises when certain norms – false and oppressive ones – impede the ability to reflect upon the norm. As a matter of contingent fact, those choices that are informed by or accord with false and oppressive norms are not autonomous. Insofar as Stoljar is aiming to sketch a substantive constraint, she presumably does not intend her claims to be understood in this way. Still, perhaps her claims here are correct, so this understanding is worth considering further.

a) A competence condition

³⁸ *Ibid*, p.108.

³⁹ *Ibid*, p.108.

⁴⁰ In chapter 3 I explore more fully what ‘choosing in accordance with’, or ‘choice inconsistent with’ might in fact involve, with respect to this part of the substantive account. I there argue that no plausible account of this can be given.

⁴¹ Stoljar, N. (2000) *Op.Cit.* p.109.

On this understanding the claim is that agents cannot competently choose in accordance with false and oppressive norms or values. When an agent accepts a false norm, the claim must go, her capacity to reflect upon the truth or falsity of that norm is undermined. Autonomy requires an adequate level of reflection that these false norms impair. Understood in this way, though, it seems false that, as a matter of empirical fact, the falsity of a norm prevents the agent from seeing that it is false, as Stoljar claims. Recall the claim, from Stoljar, that the women “accept something false. And because of the internalisation of the norm, they do not have the capacity to perceive it as false”. Suppose we understand what it is to internalise a norm in a weak sense: as the effective incorporation of a belief about what one ought to do into an agent’s belief corpus. So, the effective internalisation of the norm

(n1) women ought not be active sexual agents

will involve the effective incorporation of the belief

(b1) I (as a woman) ought not be an active sexual agent

into the agent’s belief corpus. Is there any reason to suppose that the falsity of a belief prevents agents from reflecting upon the false belief? It seems not. Some false beliefs will be easily reflected upon – and rejected. For instance, my false belief that “the man by the bus stop is Mr. Zalatimo” can be swiftly rejected, upon the acquisition of new perceptual evidence. Others may be more deeply entrenched, and hence less easily reflectively rejected. But in such cases this is not a function of falsity. If the ability to reflect on false beliefs is undermined, there is no reason to suppose that this is due to the belief’s falsity; perhaps it is rather due to its entrenchment.

Indeed, internalisation might involve something stronger than I have so far supposed, such as the entrenchment of a belief about what one ought to do in an agent’s belief corpus. The well-entrenched nature of a belief may prevent the agent from engaging in adequate reflection upon it – and this may be so with respect to beliefs about what one ought to do. But if we take this to be what presents a threat to autonomy, then many agents who act in accordance with the true and the good will fail to meet the conditions for autonomy – having well entrenched beliefs about the world and beliefs

about norms is a feature of many agents on many occasions of deliberation and choice (and surely not always an undesirable one, at that).⁴²

We should not accept Stoljar's claim that the falsity of the norm prevents the agent from adequately reflecting upon it. Nor should we reinterpret her condition as pertaining to the entrenched nature of the norm. Perhaps the problem pertains only to those norms that are both false and well entrenched? Once again, it is not clear why a norm that is well entrenched and false would prevent reflection upon it. For example, suppose that, due to limited experiences and an insular, village life, Eklas believes that women are naturally suited to domestic work, and ill-equipped to work in the public domain. Thus she accepts the false norm "women ought not work outside the home", and this is well entrenched. Then her village gets a Television set. She watches programmes that show women working in a range of roles, and starts to reflect on whether she is right to believe that women ought not work outside the home.⁴³ The falsity and entrenchment of this norm does not seem to hinder in any way her ability to reflect on this norm: indeed it may even be that its falsity helps her to reflect upon it – because she bumps up against states of affairs that conflict with the dictates of the norm, and which prompt her to reflect upon whether to accept it.

So we should reject Stoljar's claims about the falsity and entrenchment of norms undermining competence. And were we to accept these claims, then we should reject the claim for which Stoljar purports to be arguing: that only a substantive account can make sense of the feminist intuition. A competence condition, were we to accept it, is content-neutral, demanding only that the agent have the ability to reflect upon her deliberation and choice. This is compatible with the agent operating with false and entrenched norms.

b) A substantive condition

A second way to understand her account is as directly constraining the contents of autonomous choice. Stoljar envisages her proposed condition as "one which places restrictions on the contents of agents' preferences".⁴⁴ Understood in this way, the condition that she offers places a straightforward constraint on the contents of

⁴² It is this entrenchment, and the fact that we often operate on such norms unreflectively that prompts theorists to offer counterfactual conditions, such as Christman's (1991b).

⁴³ See Waldfogel, J. (2007) 'TV Is Good for You, if you are a woman in rural India, at least' at <http://www.slate.com/id/2172474/> for a report on women in rural India whose attitudes changed having been exposed to different portrayals of women's lives.

⁴⁴ *Ibid*, p.95.

autonomous choice, rather than making a claim about the impact of the norms upon an agent's capacities. We have seen Stoljar's claims that a choice is non-autonomous when it is in accordance with norms that are false and oppressive. Choices, then, must not be in accordance with false norms – conversely, they must be in accordance with true norms. Nor must an agent's choice be in accordance with oppressive norms; rather, they must be in accordance with the non-oppressive, or good. Supposing that we take 'in accordance' to mean 'not in conflict with' – then those choices that are in accordance with the neutral or non-oppressive will not conflict with the true or the good (although they may not be choices for true and good things – they may be choices for lacking-truth-value or neutral things). Such choices will be in accordance with the true and the good. So we can put Stoljar's condition in positive form, as:

(RelChoice1) a necessary condition for autonomous choice is that the agent choose in accordance with the value of the true and the good.

This condition captures the thought that, insofar as agents choose in accordance with false and oppressive norms, their choices cannot be autonomous. Note that there is another version of the condition in the offing. Sometimes Stoljar talks of norms that are false and irrelevant posing a threat to autonomy: "the fact that such norms are false or irrelevant diminish or extinguish agent's capacities for autonomy with respect to decisions governed by the norms".⁴⁵ This claim suggests the following version of the condition:

(RelChoice1*) a necessary condition for autonomous choice is that the agent choose in accordance with the value of the true and the relevant.

This version of her claim is not illuminating for present purposes. Insofar as a norm is operative in a particular context, such that it is reasonable for the agent to take it into account in deliberation, the norm is not irrelevant. In the context of deliberation about contraceptive practices, a norm that says "one ought to take shoes off when entering the house" is irrelevant. A norm that says "women ought not engage in pre-marital sex" is not irrelevant (although we may think it should not be heeded). One might think that it ought not be relevant. But if it is a norm that the agent will in fact be judged according to, and faces potential costs for violating, then it is not irrelevant to her decision making. (RelChoice1), then, is the substantive condition that I will focus on in the following.

⁴⁵ *Ibid.* p.108.

According to (RelChoice1), choices that accord with the false and oppressive cannot be autonomous. In the absence of the (competence related) claims about the impact of the internalisation of false norms upon the agent's capacities, what arguments are offered for this condition? The main claim in support of this direct constraint is the appeal to 'the feminist intuition'. This argument seems to take the form of inference to the best explanation. Stoljar has the intuition that the women are non-autonomous. What best explains this intuition – having attempted to rule out the explanatory adequacy of the content-neutral conditions – is that the women are choosing in accordance with false and oppressive norms.

i) Explanatory inadequacy

In this respect, there are three points to note. First, not all will accept the intuition, and so many will not accept the *need* for such an explanation. As we have seen, Benson notes that "feminists need not accept the intuition on which Stoljar founds her position".⁴⁶ I have, for the sake of argument, granted Stoljar this intuition. Second, some of the content-neutral conditions that Stoljar rejects may, in fact, be able to explain the non-autonomy of some of the choices – so I have argued. Third, it is not clear what explanatory power the claim "because the agent chooses in accordance with false and oppressive norms" has. The first conjunct does not seem to do much explanatory work; agents on many occasions make choices informed by false beliefs ('that accord with the false') and this does not seem to undermine the agent's autonomy of choice.⁴⁷ When an agent chooses to buy crystals to create energy flows in his house; when an agent mistakenly believes that kidney beans upset his constitution, and so chooses not to eat them; when an agent chooses to drive without a seatbelt, believing the inconvenience to be more hassle than it is worth: all these are instances in which it might well be the case that the agent acts in accordance with 'the false'; that is, on the basis of false beliefs. Yet it is far from clear that these agents are non-autonomous. In cases of failing to wear seat-belts, Stoljar herself admits as much. So choosing in accordance with the false is not intuitively autonomy undermining. Moreover, were choosing autonomously in accordance with the false impossible, then one could never autonomously make mistakes of this kind, and that seems wrong.

⁴⁶ Benson (2005b) *Op.Cit.* p.125

⁴⁷ In chapter 4 I return to, and say more about, the issue of the epistemic conditions for autonomous choice.

So, the conjunct that does the work in explaining Stoljar's intuition that such agents are non-autonomous must be that which pertains to the oppressive nature of certain norms: it is because the agent is choosing in accordance with the oppressive that she is non-autonomous. Setting aside difficulties about how one is to understand 'the oppressive', it is hard to see precisely what explanatory story could be offered here, also. One can grasp that choosing to act in accordance with the oppressive might be incompatible with flourishing agency, or with agency that promotes well-being or interests. But it is not clear how it is incompatible with autonomous choice – her ability to properly govern her deliberations and subsequent choice. For example, consider an agent who deliberates carefully about and finally chooses to take a job as a stripper in order to have enough money to feed her children. One might think that such a choice accords with oppressive values, and is incompatible (at least in current social conditions) with flourishing agency, or agency that promotes (overall) well-being. But it is not clear that this is incompatible with autonomous choice. Indeed, that the agent has deliberated carefully about what, given the difficult circumstances, the prudential option is suggests that she is perfectly able in her governance of her deliberations and subsequent choice.⁴⁸

A similar conclusion about women's choices in restrictive and oppressive social contexts is drawn by Uma Narayan. With reference to the Sufi Pirzada Muslim women who engaged in practices of purdah (seclusion) in Old Dehli, she writes that we should not understand the women who engage in these practices as forced into oppressive lives against their wills, or as blinded by culture and tradition (as 'prisoners of patriarchy' or 'dupes of patriarchy' respectively). Rather she claims they should be seen as 'bargainers': they cooperate with some cultural norms whilst acknowledging the centrality of the practices to their religion and traditions; they complain about the restrictions and inconveniences; but they also acknowledge the benefits of their roles, in terms of financial interest and social standing. That the practices are in many respects, ones that we would judge to be oppressive does not mean that the choice to comply with them is non-autonomous, she claims: 'a person's choice could be autonomous even if made under considerable social or culture pressure'.⁴⁹ Choice in accordance with oppressive norms, she claims, should not be considered non-autonomous simply

⁴⁸ One might be tempted to hold that autonomous choice is undermined in such a case due to the coercive pressures faced in such a situation. But we need not adopt a substantive condition to make sense of this problem – a historical condition will serve to make this diagnosis.

⁴⁹ Narayan, U. (2002) 'Minds of Their Own: Choices, Autonomy, Cultural Practices and Other Women', in *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, Antony, M. and Witt, C. (eds.), Boulder, Colorado: Westview Press, pp. 418-432, at p. 429.

because of the presence of oppressive norms. Absent further argumentation, and given the contested status of the feminist intuition, I suggest that we should not accept Stoljar's strong substantive condition (RelChoice1).

ii) Contrary intuitions

I have said that for the sake of argument I would grant Stoljar the intuition that the agents who act in accordance with the false and oppressive norms are non-autonomous. However, now that we have seen her condition to be unmotivated, and explanatorily inadequate, it is worth noting that there are compelling intuitions to the contrary. Diana Meyers has highlighted the capacity for autonomous choice and agency that can be exercised in conditions of oppression, and emphasised the importance of recognising this, claiming that “feminists must account for the control women exert over their lives under patriarchy, for their opposition to subordinating social norms and institutions, and for their capacity to bring about emancipatory social change”.⁵⁰ (Indeed, she suggests that those who are oppressed may well be better positioned to exercise autonomy of choice when it comes to issues of justice and policy.)

To add plausibility to Meyer's claim, consider the following example: suppose that Edna occupies a traditional role, and is dependent upon her husband for economic support. Suppose that she endorses this role. But having reflected upon her own experiences, comes to believe that, because of the position of dependency that many women occupy, it is incredibly important that women are able to control when, and on how many occasions, they have children. So Edna campaigns actively for women's reproductive rights, and policies such as access to abortion for all. Now, it may well be that in her opposition to some oppressive norms (such as “women ought not abort, it is their duty to be mothers” and so on), she makes reference to other oppressive norms, such as “primary care givers should be women”, and “women should take primary responsibility for domestic matters”. If we accept Stoljar's claim that choice in accordance with false and oppressive norms is non-autonomous, then we cannot claim that the choice of this woman, to campaign for pro-choice laws, is autonomous. This seems counter-intuitive. Despite her beliefs about women's roles, Edna seems autonomous in her choice to campaign – indeed, as the example is set up it is in part

⁵⁰ Meyers, D. (2000) *Op.Cit.* p.152. Paul Benson ((2000) 'Feeling Crazy: Self-worth and the Social Character of Responsibility' in *Relational Autonomy*, MacKenzie and Stoljar (eds.) pp.72-94) makes a similar point.

because of her beliefs about women's roles that she so actively campaigns for women's reproductive rights.

If we accept Stoljar's substantive condition, then the possibility of recognising the autonomous choices of agents in contexts of oppression is severely limited. Agents whose deliberations and choices make reference to norms that are 'false and oppressive' simply cannot be autonomous, on Stoljar's view. And this would appear to rule out from the class of autonomous many of the choices made by agents who are influenced by oppressive norms.

Meyer's claim that those in social conditions of oppression can exercise autonomy in choice is very plausible, and it seems very problematic to claim otherwise. It is a troubling consequence of Stoljar's substantive condition that this is not possible. These considerations are not conclusive. This is in part because the condition that Stoljar develops is, as she admits, underdeveloped. But I suggest that these considerations recommend against developing it further.

2.4 Summary

We have seen the first attempt to offer a condition for autonomous choice that attempts to take into account the relational aspects of autonomy, by building in substantive conditions. This account explicitly aims to make discriminations, in accordance with the 'feminist intuition', in a way that makes sense of those agents whose choices, it is claimed are non-autonomous having been negatively affected by gendered social relations.

I have argued that we should reject Stoljar's claims; first she does not do enough to show that content neutral conceptions cannot diagnose those whom she wishes to as non-autonomous; second, the substantive conception she proposes is explanatorily inadequate, as it is not clear what explanatory work can be done with the claim that it is because of the accordance of the choice with values that are false and oppressive that these women's choices are non-autonomous. Third, there are good reasons to reject the intuitions about non-autonomous choice that motivate Stoljar's account.⁵¹ Indeed, the contrary intuition garners support from Meyers' claim that many

⁵¹ And of course, if the intuition is rejected, it will not be seen as a failure of content-neutral conditions that they fail to make sense of it.

in oppressive social contexts exercise significant self-governance, not least in their attempts to bring about social change.

It is worth noting, however, that whilst we might reject the intuition that such agents are non-autonomous, it seems that there is nonetheless something troubling about choices that are influenced by norms which are, at least from a feminist point of view (although I hope beyond, also), criticisable. Accepting that the problem is not a matter of non-autonomous choice does not mean that other diagnoses or criticisms cannot be offered, although I shall not attempt to do so here. But there is more yet to be said about substantive conditions. One might, for instance, incorporate a different value into the substantive constraint on autonomous choice. I turn to consider a view which posits self-respect as the constraining value on autonomous choice.

Chapter 3. Substantive account II: Hill Jnr., Baron

We have seen one attempt to incorporate into a relational condition into a conception of autonomous choice, from Stoljar. I here look at another, assessing whether it fares better. Like Stoljar's the condition offered by Thomas Hill Jnr. is an attempt to substantively constrain the contents of autonomous choice.¹ This yields a conception of autonomy that can purportedly make sense of why certain social relations are incompatible with autonomy.

The choice for a deferential role has been taken, by a range of theorists, to be paradigmatically non-autonomous.² Feminist philosophers too have found women's choice for and occupancy of the role of the deferential wife troubling.³ Thus, if we are to determine a plausible substantive constraint on the objects of autonomous choice, a look at the characteristics of the choice for deference seems to be a good place to start. I here look at one prominent account, according to which, first, such choices are judged to be non-autonomous, and second, the non-autonomy of these choices is diagnosed and explained by a substantive conception of autonomy. This substantive condition is specified as:

(RelChoice2) a necessary condition for autonomous choice is that the choice is in accordance with the value of self-respect.

Now, the theorists who propound and defend this condition do so in a context that precedes the recent debate about the nature of 'relational autonomy'. The main motivation is to make sense of what is deemed to be intuitively problematic about these

¹ See Hill Jnr., T. (1991) 'Servility and Self-Respect' in his *Autonomy and Self Respect*, Cambridge, Cambridge University Press, pp.4-18. Originally printed in Wasserstrom, R.A. (ed.) (1979) *Today's Moral Problems*, 2nd Edition, pp.133-147. (References are to the 1991 version.) His view is defended against criticism from Friedman, M. ((1985) 'Moral Integrity and the Deferential Wife' *Philosophical Studies* 47 (1), pp.141-150), by Marcia Baron (1985), 'Servility, Critical Deference and the Deferential Wife' *Philosophical Studies*, 48 (3), pp.393-400.

² This is not uncontroversial: a range of theorists – those with content-neutral accounts – maintain that deferential roles, such as slavery, can be chosen autonomously. See Dworkin (1988) *Op.Cit.* for such claims. Again, for the sake of argument I initially grant the intuition that such agents are non-autonomous.

³ See Westlund, A. (2003) *Op.Cit.* See also Superson, A. (2005) 'Deformed Desires and the Informed Desire Test' *Hypatia*, 20 (4) pp.109-126.

choices, and to distinguish them from choices that are not, intuitively, problematic. However, the focus of Hill's original article can plausibly be understood as an attempt to make sense of how social conditions, namely, conditions of social disadvantage, might be non-conducive to the kind of self-respect that is necessary for autonomous choice.⁴

In this chapter, I explore whether the substantive condition proposed is able to make sense of the reported intuitions, and thus diagnose the problem with the deferential individual. I will argue that it cannot. Any support for a substantive condition that such a diagnosis might have offered, then, cannot be relied upon.

3.1 The problem with deference

The diagnosis of Thomas Hill Jr. and (in defence of Hill) Marcia Baron, of the problem with a choice for deference, is Kantian in nature. Thus it has a distinctive moral tinge to it. The failing is understood as a failing of autonomy; but proper self-governance, in Kantian terms, requires that an agent govern herself in accordance with the moral law. It is no surprise, then, that such a conception of autonomy will incorporate substantive content. Indeed, at this level of generality the substantive constraint can be formulated as:

(RelChoice2*) a necessary condition for autonomous choice is that the choice is in accordance with (the value of) the moral law.

Unpacking what the moral law requires is no small task. Here, I focus on just one aspect of the substantive constraint that may be entailed by the moral law – that intended to do the work in diagnosing the defect of the deferential individual: namely, the requirement that autonomous choice is in accordance with the value of self-respect. For an agent to choose autonomously, she must do so in accordance with the value of respect for humanity, and this “amounts to having a kind of self-respect [which is] incompatible with servility”.⁵

⁴ See Hill, T. (1991) *Op.Cit.* pp.5-7.

⁵ Hill, T. (1991), *Op.Cit.*p.13. Note that this strategy – of setting aside the specifics of the moral law – is consistent with Hill's. As Susan Mendus' review observes, Hill's “essays do not engage in detailed analysis of Kant's texts. Rather, Hill begins with some everyday problems of moral life, and shows how, in order to solve these problems, we must appeal to Kantian insights”. See Mendus, S. (1992) Review: *Autonomy and Self-Respect*, Hill Jr., T. *Philosophy*, 67 (262) pp.561-563. My strategy here is to show that Hill's Kantian insights cannot help in diagnosing the problem with deferential choices, at least.

What is at stake here is not deferential behaviour simpliciter: there might sometimes be reasons, consistent with self-respect, for behaving deferentially (prudential reasons, epistemic reasons, and so forth). It is being deferential, or servile, in attitude that is problematic. An example will help to clarify the kind of deference that is deemed to be problematic (although as we shall see, spelling out more precisely what this attitude consists in is extremely difficult):

Deference: having completed her university degree in maths, Rose is deliberating about whether to pursue further research, or whether to get married to Chad and take on the role of housewife. Marriage would mean turning her back on further education and a career, and would require her to put all her energies into looking after her husband, maintaining the household, and bringing up their children. Both Rose and Chad have had upbringings, and live in communities in which typically, the man makes the important decisions for the family, and the woman has little input other than in the daily execution of the husband's legislation. Both believe that this is a good way to structure a family. After some careful deliberation, in consultation with Chad, but also independently, Rose decides that she would like to settle down and get married. From then on, Rose is "utterly devoted to serving her husband. She buys the clothes *he* prefers, invites the guests *he* wants to entertain, makes love whenever *he* is in the mood ... she is quite glad and proud, she says, to serve her husband as she does."⁶

Making such a choice, and occupying such a role, is not consistent with the kind of self-respect required for autonomy, Hill and Baron both claim. Let us suppose that Rose meets all of the content-neutral requirements that might be demanded (an overview of which we saw in the last chapter): she endorses her choice; she formed her preference in a way that she would not resist; she does not have inconsistent beliefs or desires that lead to a breakdown of critical faculties; her reflective powers are not inhibited; she has a normal level of self-knowledge, and is free from coercion and manipulation when she makes her decision. If she is non-autonomous in so choosing, then, it must be due to something else – perhaps due to the substantive content of her choice.

This is the contention of the Kantian substantive view, as defended by Hill and Baron. The servile person's choices fail to be autonomous, on this view, because they fail to meet the following necessary condition:

(RelChoice2) a necessary condition for autonomous choice is that the choice is in accordance with the value of self-respect.

⁶ *Ibid*, pp. 5-6.

The claim is that this condition is derivable from the moral law: it is entailed by the necessary condition (RelChoice2*). The moral law demands that agents must possess “a certain attitude concerning one’s rightful place in a moral community”.⁷ This attitude, in particular, is one according to “which [s]he is equal with every other person”.⁸ That is to say: all persons must be respected as equal in their moral status; so the moral law demands. The ‘all’ here applies universally – so in addition to respecting others, an agent must also respect herself, and see herself as equal, with regards her moral standing, to others.

It might appear that with this condition we can already diagnose the problem with deference – for it is intuitive that the deferential agent such as Rose does not see herself as equal to others – in particular, with those to whom she defers. But we cannot yet reach this conclusion: the dimension along which one must take oneself to be equal is that of moral status. Whilst the deferential person might not see herself as a social equal, or equal in terms of expertise, say, why should we think that the deferential person fails to treat herself as a moral equal, as failing to treat herself as occupying her ‘rightful place’ amongst moral agents? In order to see this, we need to ask: ‘What does self-respect involve’?

3.1.1 Two ways of lacking self-respect

The Kantian view provides us with the following answer. Having the requisite self-respect consists in according:

SUFFICIENT WEIGHT TO SELF: S must accord appropriate weight to her interests, preferences, or rights.

Now, one way to understand the claim that one must give appropriate weight to one’s rights or preferences is as a procedural matter: one must have true beliefs about the value of one’s preferences, and in the process of deciding what to do, one must not violate norms of deliberative rationality failing to weigh one’s rights or preferences in accordance with these beliefs about their value. Understood thus, the self-respect requirement does not amount to a substantive constraint, and the account collapses back into a procedural, content-neutral account.

⁷ *Ibid*, p.6.

⁸ *Ibid*, p.9.

But this does not seem to be the way that Hill or Baron intend the account, and we should not understand it in this way. Rather, the claim is that agents have a certain value, and from this value is generated the value that attaches to their interests, preferences and rights. An agent's choice must be in accordance with the value of her agency and this means not allowing others to trample on her rights or override her preferences. Insofar as acting deferentially involves this, it is not consistent with the value of agency, of self-respect, or of the rights and preferences of the agent.

That is, having the kind of self-respect demanded by the moral law requires that one choose in accordance with the value of one's agency – one's interests, preferences or rights. A constraint on autonomous choice is that one must choose in accordance with this value. There are two glosses on this condition. According to Hill, the deferential agent fails to meet (RelChoice2) because she is guilty of “a failure to understand and acknowledge [her] own moral rights”.⁹ The deferential person “tends to deny or disavow [her] own moral rights because [she] does not understand them or has little concern for the status they give [her]”.¹⁰ In choosing deference, and subsequently being deferential, Rose fails to choose in accordance with the value of her rights. In so failing, she does not display the requisite self-respect. Thus Hill's specification of the substantive condition is:

(RelChoice2a) a necessary condition for autonomous choice is that the choice is in accordance with the value of her moral rights.

Baron, in her defence of Hill's claims, focuses rather on the other disjunct of SUFFICIENT WEIGHT TO SELF, such that the specification of the condition is:

(RelChoice2b) a necessary condition for autonomous choice is that the choice is in accordance with the value of her interests or preferences.

In wearing what her husband prefers, rather than what she prefers, Rose fails to choose in accordance with (the value of) her preferences, for example. This is a failing of self-respect. In the following, I consider whether either of these conditions is able to do the work required of it, in terms of capturing the intuitive non-autonomy of her deferential choices.

⁹ *Ibid*, p.9.

¹⁰ *Ibid*, p.13.

3.2 Moral rights

Hill's claim is that the individual who chooses deference does not properly understand – or care for – her moral rights, namely, her moral right not to defer. If this is a moral right, then being treated as deferential may constitute a rights violation, and one inconsistent with respect for the agent. Does this mean that in taking on such a role, and allowing others to treat her as deferential, an agent lacks self-respect? This is not clear: an agent can choose to waive her rights, or not to exercise them. Why should the deferential individual be seen as misunderstanding her rights, rather than legitimately waiving them? If we are to accept Hill's claim that in choosing deference an agent fails to respect her moral rights, and hence fails to meet (RelChoice2), we need to see, first, that the right not to defer is derivable from the moral law, and second, that such a right is unwaivable.

It is claimed that the right not to defer is “derivable from one's duty to respect the moral law”,¹¹ but the argument for this is not explicit. It can be reconstructed as:

Argument 1: an argument for the right not to defer

1. Treating others as servile, or deferential, fails to treat them with respect
2. The moral law demands that one treats all persons with respect – persons have a duty to treat others with respect.
3. Persons have a duty not to treat others as servile, or deferential
4. Each person has a right, corresponding to this duty, not to be treated as servile, or deferential.

This argument establishes that each has the right not to be treated as deferential. And if one has a right not to be treated as deferential, then one has a right not to act deferentially, or occupy a role of deference; not to defer (compare: if I have a right not to be treated as team leader, I have a right not to act as leader, or occupy the team leader role: that is, a right not to lead).

This may be, then, how right not to defer is derivable from the moral law. But for deferring to be a failure of self-respect, we need to see that this is a right that cannot be waived. If an agent can waive the right to defer, then she can remove from (certain) others the duty not to treat her as deferential. In order to see that allowing others to treat her in this way is a failing of self-respect – a case of allowing others to ‘trample on her

¹¹ Baron, M. (1985) *Op.Cit.*, p.394.

rights', as Hill puts it – we need to see either that this right is unwaivable, or that there is some intrinsic moral failing in allowing herself to be so treated. The following argument is reconstructed from Hill (see pp.9-14):

Argument 2: first argument for the right not to defer being unwaivable

1. Each person ought to respect the moral law
2. Respect for the moral law requires not only acting in accordance with the moral law, but also respecting – treating as valuable – “the provisions of morality”.¹²
3. The provisions of morality include the principles, ideals, rights and duties derivable from the moral law.
4. So each person ought to respect and value the rights derivable from the moral law, including the right not to defer.
5. So each person ought to respect and value the right not to defer, by not allowing others to trample on this right, and by not waiving this right.

The thought here is that to waive the right not to defer is not to properly value it. In failing to properly value her right, an agent who waives it fails to properly value herself – namely, she lacks self-respect, so fails to meet (RelChoice2). However, argument 2 on its own cannot establish this. We need an independent argument to see that all rights derivable from the moral law are unwaivable. Otherwise, it is quite consistent with valuing one's rights that one waives it. For example, suppose I have a right that others not inflict pain on me. I can waive this right, permitting others to inflict pain on me; perhaps in undergoing an endurance test for medical research, or in having a tattoo. Waiving this right, on these occasions, is consistent with me valuing very much and quite properly the right that others not inflict pain on me.

The rights that cannot be waived, Hill tells us are those that are grounded in an agent's “inner freedom”,¹³ that is, grounded in the capacity for autonomy that individuals have. But the point here cannot amount to the claim that the right not to defer is unwaivable because so deferring involves a failure of self respect, and so non-autonomy in choice. This is precisely what argument 1 & argument 2 have been trying to show. If the arguments for the conclusion that deference and autonomy are incompatible rely on the presupposition of this very claim, then they will be problematically question-begging. Independent support for this claim about the incompatibility is required, if it is

¹² Hill, T. (1991) *Op.Cit*, p.13.

¹³ Kant, I. *The Metaphysics of Morals*, (1797/1996) Gregor, M. (ed.), Cambridge, Cambridge University Press, see esp. [6:407-6:408], at pp. 165-166.

to support the claim that the right not to defer falls into the class of rights that cannot be waived.

A further argument can be garnered from Hill's comment that "a man who consents to be enslaved ... thereby displays a condition of slavish mentality that renders his consent worthless".¹⁴ This argument is made explicit by Baron:¹⁵

Argument 3: second argument for the right not to defer being unwaivable

1. failing to properly understand one's moral status – including the value of one's moral rights – indicates lack of autonomy.
2. consent that is non-autonomously given is not valid consent.
3. if an agent e.g. treats herself as deferential, allowing others to treat her as deferential by waiving her right not to defer, then an agent does not properly understand her moral status.
4. so if an agent treats herself as deferential, allowing others to treat her as deferential by waiving her right not to defer, then she is non-autonomous in doing so.
5. so an instance on which an agent waives her right not to defer and allows others to treat her as deferential is an instance in which the agent's consent is not valid.
6. rights can only be waived with valid consent.
7. the right not to defer cannot be waived.

If valid, this argument could establish that the right not to defer cannot be waived, and that any attempt to do so is a failure of self-respect such that (RelChoice2) is not met. We would have a condition that could make sense of the intuitive non-autonomy of the choice for deference. But this argument too is question-begging. Premise 3 of argument 3 is:

3. if an agent treats waives her right not to defer, then an agent does not properly understand her moral status

But if the justification for this claim is the thought that an individual who did understand her rights properly would recognise that the right not to defer is unwaivable,

¹⁴ Hill, T. (1991) *Op.Cit.* p.15.

¹⁵ Baron, M. (1985) *Op.Cit.* pp.396-397.

the premise is assuming what the argument is aiming to establish. Hill's arguments, then, appear to be beset by circular reasoning. We cannot find a non-circular grounding for the claim that the deferential individual fails to respect her moral rights. A diagnosis of Rose's choice as non-autonomous cannot be grounded in the claim that she fails to respect her moral rights, and hence lacks self-respect.

An alternative way to make sense of the claim that the choice for deference is incompatible with the value of self-respect might be found in Baron's suggested interpretation: that the individual who chooses a deferential role fails to accord appropriate weight to her interests or preferences. I turn to this interpretation now.

3.3 Interests or preferences

Baron's understanding of the constraint on autonomous choice, recall, takes the following form:

(RelChoice2b) a necessary condition for autonomous choice is that the choice is in accordance with the value of her interests or preferences.

There is an intuitive sense of what it is to fail to choose in accordance with one's preferences. This is what Baron has in mind with her example of a wife, such as Rose, who gives insufficient weight for her preference not to wear make-up:

we say to her, 'oh, so you prefer wearing make-up to not wearing it?' 'No', she replies, 'I'd rather not wear it; but my husband would hate that.'¹⁶

Here I want to examine what substance we might give to this intuition, and whether we can make sense of the way in which the deferential individual might fail to give appropriate weight to her preferences such that she fails to meet (RelChoice2b). If good sense can be made of this claim, then the following argument could be made:

Argument 4: for self-respect requiring appropriate weight to preferences

1. An agent treats herself with self-respect, in accordance with her status as a moral equal with others, if she treats her preferences as having (prima facie) equal weight with the preferences of others¹⁷

¹⁶ Baron, M. (1985) *Op. Cit* p.395.

2. the moral law demands that all persons are treated as moral equals – that all persons treat everyone (including themselves) with respect
3. if an agent is deferential, she fails to treat herself as having equal moral status – because she takes her interests and preferences to have less (prima facie) weight than those of others.
4. if an agent is deferential, then, she fails to treat herself with self-respect, in accordance with her status as a moral equal with others

If this argument can establish that agents who are deferential fail to properly recognise their moral status argument 3 can then run, utilising the claim that waiving the right to defer cannot be autonomously performed. Or a more direct argument could be constructed: if an agent attempts to waive the right, she does not properly understand her moral status. Failing to understand that this status and the rights that it entails is a failure of self-respect. Hence an agent who chooses deferential roles fails to meet (RelChoice2). However, I will suggest that it is very difficult to make sense of the seemingly intuitive thought that might support premise 3 of argument 4.

3.3.1 Intrinsic/first order desires

A first attempt to make sense of the intuitive claim might state that an individual who chooses a deferential role fails to give adequate weight to her own desires: in that role her choices and actions will primarily be geared towards the satisfaction of the desires of others. This thought is challenged by Marilyn Friedman, who writes that:

the fact that a person takes, as the preferences which she will act to satisfy, preferences which happen (originally) to be those of other persons, does not entail that she acts on preferences which are not (in the last analysis) her own. ... [She] takes the preferences of certain other persons for her own; they become her preferences.¹⁸

The thought here is that the deferential individual may nonetheless act on preferences that are – having become – her own. Baron finds this claim wrong-headed, responding that;

¹⁷ The qualifier ‘prima facie’ is required here, because one might, upon reflection decide that someone else’s preference to spend their spare time saving dolphins is more weighty than one’s own preference to spend one’s spare time watching Big Brother.

¹⁸ Friedman, M (1985) *Op. Cit* p.144.

to say that A accepts a view which leads her or him (cheerfully) to abide by someone else's wishes (preferences) is not to say that A adopts this person's preferences. Of course A might come to adopt the other's preferences; but it is quite possible – and extremely common – for such a person not to.¹⁹

One way to understand the claim here is that the deferential individual fails to give weight to her intrinsic desires. She chooses in accordance with her husband's preferences – say, wearing a red shirt, when she has an intrinsic desire for a blue shirt. Her desire for wearing the red shirt is merely instrumental: instrumental to the satisfaction of her desire to please her husband. Thus she fails to accord appropriate weight to her intrinsic desire (for the blue shirt). But it is difficult to see what is problematic with choices that have this structure. I have an intrinsic desire to sit in the sun and eat mango. My desire to stay inside and write is instrumental – instrumental to the satisfaction of my desire to write a chapter that will be part of an interesting thesis, which is instrumental to the satisfaction of my desire to be a good candidate on the job market... and so on. I forsake sitting out eating mangoes for the sake of the instrumental desire to write my thesis. This is quite normal, and there is nothing troubling with a structure of preference and choice in which an agent forsakes the satisfaction of an intrinsic desire for an instrumental one.

Alternatively, Baron's thought might be understood as the claim that the deferential individual fails to give appropriate weight to her first order desires. However, on one widely accepted view of moral psychology, when an agent chooses and acts, there is, of necessity, a motivating desire on which she acts. On this view, the standard Humean view, if an agent performs action x, she has a desire to do x.²⁰ Such a claim about motivating desires is not the solely held by Humeans: it is also accepted by prominent Kantians. Korsgaard, for instance, writes that:

Kant thought that every action involves some incentive [desire] or other, for there must always be something that prompted you to consider the action²¹

¹⁹ Baron, M. (1985) *Op.Cit* p.396.

²⁰ This view has its roots, naturally, in Hume, who claimed that "reason alone can never be a motive to any action of the will... Abstract or demonstrative reasoning, therefore, never influences any of our actions, but only as it directs our judgment concerning causes and effects" (*A Treatise of Human Nature* ((1888/1978) Nidditch, P.H. (ed.), 2nd Edition, Oxford, Clarendon Press) Book II Part III at pp. 413-414). If an agent is to be moved to action, then, a passion – or desire – must be present to give the motivational push. Thus we have the 'belief-desire pair' view of action, that informs much moral psychology, as well as the 'folk psychological' view of human action.

²¹ Korsgaard, C. (2002) 'Practical Reason and the Unity of the Will' *Locke Lecture II*, p.24. Available online at: <http://www.people.fas.harvard.edu/~korsgaard/#Locke%20Lectures>.

and:

every action must involve both an incentive [desire] and a principle²²

On this view, when the deferential individual acts, even in cases where she chooses to act deferentially, she does act on a first order desire or preference that is her own. Thus she gives weight to her first order desires. When Rose buys a red shirt rather than a blue one, because Chad prefers red to blue, she gives weight to her first order desire for a red shirt (and does not give weight to her first order desire for a blue shirt).²³ So this cannot be the right way of making intuitive sense of the thought that such an individual fails to give appropriate weight to her preferences.

3.3.2 Higher order desires

A natural progression, then, would be to look to the agent's attitudes toward her first order desires; her reflective endorsements or rejections of the desires that move her to action (a view we encountered in the previous Chapter, in assessing Frankfurt's account of higher order endorsements). Perhaps the problem with the deferential wife is that she fails to accord sufficient weight to what she really wants, namely, those wants that are supported by her reflective endorsements. On this view, whilst the deferential individual may want to do x (where x is determined by the individual to whom she defers), she does not want to want to do x.²⁴ Her first order desire lacks the kind of supporting higher order attitude (desire) that can serve to turn her run of the mill first order desire into a preference in the sense required for her to meet (RelChoice2b). On this view, preferences are the desires behind which an agent stands (with her higher order desire), as it were. Self-respect requires that an agent must give weight to her preferences in this sense.

Again, this is an intuitive understanding of the notion of a preference, but not one that could do the work in establishing the claim that the problem with the individual who chooses deference fails to accord her preferences adequate weight. It is plausible to suppose that the deferential individual wants to want to do what the person to whom she defers legislates: if she is to act in accordance with the dictates of another, she must

²² Korsgaard, C. (2002) 'Autonomy, Efficacy and Agency' *Locke Lecture III*, p.23. Available online at: <http://www.people.fas.harvard.edu/~korsgaard/#Locke%20Lectures>.

²³ Recall, from chapter 2, that first order desires are understood dispositionally: thus it makes sense to attribute desires to Rose for both, even if she feels a phenomenological urge to buy a blue shirt.

²⁴ See Frankfurt, H. (1971) *Op. Cit.* p.9.

surely want to have the motivating desire that can move her to so act. For instance, Rose has first order desires for both red and blue shirts. Which of these first order desires does she want to move her? She endorses the desire for the red shirt, because she knows her husband will be happy if she buys the shirt he prefers, and this is her main goal. If Rose fails to give appropriate weight to her preferences in such choices, this cannot amount to a failure to give appropriate weight to those desires that are endorsed at a higher order. Likewise, insofar as the agent has a higher order desire for the deferential role in making her initial choice – as Rose does, we assumed – she does not fail to give weight to her preferences in this sense.

3.3.3 Evaluative endorsements

The notion of ‘a preference’ seems to contain some evaluative content; the agent prefers x when she is motivated towards x and when she takes x to be valuable or important to her. The fault of the deferential individual, then, might be that she fails to give appropriate weight to she takes to be valuable. Her preferences consist in a subset of her desires; those that accord with what she values.²⁵ She takes y to be important, but fails to give weight to this – fails to choose or act in accordance with it. Rather, she chooses and acts in accordance with z, which is what the person to whom she defers prefers.

This might be true of some deferential individuals. But surely there will be some individuals for whom we cannot accept this as a diagnosis of her failing to give appropriate weight to her preferences. If Rose considered a career in maths important and failed to give weight to this, we could say she fails to give weight to her preferences. But it is entirely plausible that what an agent takes to be important is the role of serving her husband. Insofar as Rose values such a role itself, she will count as giving appropriate weight to her preferences. Such an individual will meet (RelChoice2b). The substantive theorist presumably wants to make sense of the non-autonomy of all of those with a deferential attitude: those who evaluatively endorse it as well as those who do not.²⁶

²⁵ See Watson, G. (1975) *Op.Cit* for a view according to which the evaluative stance of the agent is important in determining ‘where the agent stands’.

²⁶ As I have already noted, some theorists – Westlund, A. (2003) *Op.Cit*, Oshana, M. (2006) *Op.Cit* – maintain that an agent who fails to be wholehearted about her deference, who has preferences for other options that go unsatisfied, strikes us as intuitively more autonomous than an agent who is wholehearted in her deferential role, and fails even to form preferences of her own.

3.3.4 A generalising strategy:

It is difficult, then, to make sense of the way in which it is true of the deferential individual that she ‘fails to give appropriate weight to her preferences’. The possibilities that I have run through are not, of course, exhaustive. But I take them to be illustrative of the way in which any such attempt will fail to make good sense of the claim: insofar as the agent chooses such a role, and in her subsequent deferential actions, she will have a desire, or appropriate motivating state, to do so. And insofar as she has reflectively chosen to act in a deferential manner, or to occupy a deferential role, because she takes this to be important or valuable, then she will be giving weight to her desires or evaluative judgments.

These two features of an individual who chooses deference mean that diagnoses of any sense in which she fails to give weight to her preferences will have to spell out why it is that these preferences are not the right ones to which she should be giving weight. But this is to make a substantive judgment about what can be the object of autonomous choice, and the claim that the agent fails to give weight to her preferences is itself supposed to be what makes sense of the substantive constraint on autonomy: that the agent who chooses deferentially lacks self-respect, and, in failing to choose in accordance with this value, lacks autonomy. The prospects for this strategy of making sense of (RelChoice2), then, do not look good.²⁷

3.3.5 Preference formation

A different tack is suggested by Baron’s claim that part of the problem with a deferential individual is that she “may cease to develop her own interests”.²⁸ This might involve failing to form preferences of one’s own, or failing to sustain and foster one’s own preferences. In line with this thought, Baron imagines:

²⁷ In his ‘Self-respect Reconsidered’, Hill suggests that there is another way in which one might fail to manifest self-respect: by having personal standards, and failing to live up to them. (One might also fail to have such standards at all, which Hill too takes to be a failing). By personal standards, Hill means a set of values that one wishes to live one’s life by, whilst recognising that others may not and need not accept these particular values. Might this notion of self-respect help in diagnosing the problem with the deferential individual? It seems not: first, Hill makes clear that he intends his analysis that focuses on moral rights to remain in place as diagnostic of the failing of servility. Second, insofar as the personal standards “form of self-respect would require that one develop and live by a set of personal standards by which one is prepared to judge oneself” it is clear that deferential individuals – such as the subservient housewife – may have and live up to these standards. See Hill, T. (1991) ‘Self-respect Reconsidered’ in his *Autonomy and Self Respect*, p.22

²⁸ Baron, M. (1985) *Op.Cit* p.396.

suppose that I loved singing in choruses, but gave it up when I married because my spouse doesn't like them ... or wants me to be at home in the evenings, or because in my town the chorus rehearses on the same evening that he wants us both to play bridge. Suppose too that I no longer wear some of my favourite clothing since learning that my spouse thought it ugly. And so on. Something seems wrong.²⁹

Perhaps, then, the agent fails to give weight to her preferences by failing to sustain and foster her own preferences, as in this example. This, too, is an intuitive thought, and one that might be given substance if we consider higher order policies: policies that agents might have about the role that their desires play in deliberation, or about the weighting of desires, or about the formation of intentions, or about the formation of desires. The deferential individual may have a policy which claims:

(P) endorse first order desire for x unless my husband wants not-x.

Or,

(P*) endorse first order desire for x only if my husband wants x.

Policy (P) does not seem particularly problematic: we often form, or give weight to desires on the basis of a conditional policy. For instance, Rosalind Hursthouse and other virtue ethicists find intuitive the following policy:³⁰

(P1) endorse desire for x (unconditional honesty in action) unless the virtuous person wants not-x (conditional honesty in action).

Or, more mundanely:

(P2) endorse desire for x (parsnip soup) unless prospective guest wants not-x (not parsnip soup)

But we tend to find individuals who only sustain and develop their preferences in alignment with what others want disturbingly lacking in a sense of identity.³¹ However, this kind of consideration is quite a different kind of worry from that expressed by (RelChoice2b). The worry here is not that the deferential individual gives insufficient

²⁹ *Ibid*, p.398.

³⁰ See Hursthouse, R. (1999) *On Virtue Ethics*, Oxford, Oxford University Press. Esp. chapter 6 and 7

³¹ Although note that the policy '(P1*) endorse desire for x only if the virtuous person wants x' does not seem so troubling either.

weight to her preferences, but rather that she fails to sustain her own preferences. But failing to sustain preferences cannot simply be understood as a matter of failing to give weight to one's preferences: agents frequently alter their preferences (I haven't sustained my preference for wine; I have sustained my preference for papaya). Often this is a laudable thing (many try hard to eliminate their preferences for smoking). So altering or 'dropping' preferences, or failing to sustain some of one's own preferences, is not in itself troubling.³²

3.3.6 Doing without desires

The views that I have considered so far have utilised a moral psychology that presupposes that if an agent acts to do x, then she has a desire to do x. This moral psychology fits most naturally with Baron's expression of the concern that she fails to give adequate weight to her preferences. However, it is also worth briefly considering if better sense could be made of the claim that the deferential individual fails to give weight to her preferences on a view that maintains a different basic moral psychology. On Scanlon's view, desires need not motivate. Rather, seeing r as a reason for doing x can be sufficient to move an agent to act. On this view, then:

it is not the case that whenever a person is moved to act he or she has a desire... [W]hat supplies the motive for this action is the agent's perception of some consideration as a reason, not some additional element of 'desire'.³³

This provides us with the following way of making sense of the problem with the deferential individual: perhaps such an agent fails to give weight to the considerations that she takes to be reasons to act.

So failing, however, is not in itself normatively significant: it is part of the process of reasoning and deliberating that, upon reflection, we do not attribute weight to some of the considerations that we initially held to be reasons to act. We must

³² Might a Kantian hold that an agent has a duty to form preferences? Perhaps, but two points are in order: first, understood one way, as demanding that we hone and develop our preferences, this must be a wide duty – failing to fulfil it is akin to failing to (e.g.) be as generous as one might be. This is not the kind of fault that Hill and Baron find with the deferential individual. Second: alternatively, the duty may be understood as to require that the agent form desires. The deferential individual does this; failing to do so would surely amount to suicide. Or, this line would require something to be said about what constitutes an adequate basis for forming one's own preferences. It is not clear what this basis would amount to.

³³ Scanlon, T.M. (1998) *What We Owe to Each Other*. Mass. Harvard University Press. p.39-41.

understand the problem, if it is one, as the deferential individual failing to give weight to the considerations that she takes all things considered to be reasons to act.

Understood thus, though, we lack the ability to make any diagnosis of the problem. Suppose that Rose holds that her husband's preferences or command or request that she do x provides a reason to do x; and, when weighed with the other reasons in the balance, that her husband's preference gives her, all things considered, a reason to do x. In such a case she does not fail to give weight to the considerations that she takes to be, all things considered, reasons for action. We might criticise what she takes to be reasons for acting. But we have not yet seen an argument for the claim that choosing for bad reasons is incompatible with choosing autonomously.

3.3.7 Anticipating epistemic issues

I noted that some cases of deference seem unproblematic: choosing, or forming preferences in accordance with the dictates of the virtuous person does not seem troubling in the way that an individual who defers, say, to her husband does. It is tempting to think that the problem with the deferential individual in the latter case is that she has false beliefs about the authority of the person to whom she defers (rather than being anything to do with her desires or preferences). We defer to the proclamations of scientists, historians, grammarians, and so on, with relative frequency. One might think that what justified deference in these kinds of cases is that the person to whom we defer is a legitimate authority on the matter.³⁴ In cases in which an individual chooses a role in which she is deferential to her husband, we are inclined to think that the agent has false beliefs about the authority of the person to whom she defers.

If this thought is to make sense of the problem with the deferential individual, then an account of autonomous choice would have to incorporate:

(E) a necessary condition for autonomous choice is that the agent has true beliefs about the legitimate authority of any person to whom she defers in making her choice

My response to this is in the form of a promissory note that anticipates an argument that appears in the next chapter, to the effect that there are reasons to doubt that there are such epistemic conditions for autonomous choice. It is worth, at this stage, noting the

³⁴ See Dworkin, G. (1988) *Op. Cit* esp. chapters 3 & 4.

following two points: first, (E) looks implausible. Lily the geneticist wants to construct a good methodology for cloning in the lab. She believes Professor Whang Woo-Suk to be an authority on such matters, and having read his papers on the cloning advances he has made, she incorporates aspects of his methodology in her lab work. But her beliefs about his authority are false: his claims about cloning success are fraudulent.³⁵ This does not seem to undermine her autonomy – no less her self-respect – in her choices of experiment structure. So (E) does not have intuitive plausibility. Second, in any case, such a condition will not help Baron in her argument for the condition (RelChoice2b); it is not concerned with the weight that the agent gives to her preferences or interests.³⁶

3.4 Summary

In this chapter, we have seen another attempt to substantively constrain the contents of autonomous choice. This attempt is primarily motivated to make sense of intuitions about the non-autonomous choices of the deferential individual. However, in this chapter, I have argued that it is very difficult to give any substance to the seemingly plausible claim that deferential individuals lack the self-respect required for autonomy. I argued that Hill Jnr.'s argument was problematically circular, and Baron's defence of his claims avoided circularity at the cost of plausibility: the claim that the deferential individual fails to give weight to her preferences could not be substantiated. Incorporating a condition such as:

(RelChoice2) a necessary condition for autonomous choice is that the choice is in accordance with the value of self-respect

into a conception of autonomous agency, then, cannot do the work required of it in terms of capturing intuitive distinctions between autonomous and non-autonomous agents.

As with the discussion of Stoljar's substantive condition, though we are left with an intuition that appears in need of explanation: that there is something troubling

³⁵ See Connor, S. (2006) 'Scientist's Cloning claims Untrue ... Except for the Dog' *Independent on Sunday* Jan, 11 2006, accessed 11/05/08
<http://www.independent.co.uk/news/science/scientists-cloning-claims-untrue-except-for-the-dog-522484.html> .

³⁶ It might be a matter of failing to accord oneself sufficient credibility, however. Miranda Fricker (in her (2007) *Epistemic Injustice: Power and the Ethics of Knowing*, Cambridge, Cambridge University Press) discusses at length the failings of according others lower epistemic credibility. She does not, however, explore in detail the case of according oneself low credibility: this seems to be an interesting possibility to consider further.

about Rose's choice for a deferential role, although this may not amount to being a problem of lacking autonomy in choice. Other values – equality, justice – may be appealed to in explaining the intuition that there is something troubling here, but I cannot undertake this task here. At this stage, we can conclude that neither of the two relational conditions thus considered – both of which have taken the form of a substantive constraint on autonomous choice – can draw the intended distinctions between intuitively autonomous and intuitively non-autonomous choices. In the next chapter, I argue that no condition that has this structure can do so.

Chapter 4. A schematic argument against substantive accounts

In the last two chapters, I looked at two substantive relational conditions for autonomy. These conditions constrain the scope of autonomous choice, demanding that the agent choose in accordance with a certain value if she is to meet the conditions for autonomy. These accounts provide one way of building relational content into a conception of autonomy: certain social relations or options that are inconsistent with that value cannot be autonomously chosen (for instance, those that involve deferential roles inconsistent with self-respect); or, social conditions that fail to foster appreciation of the value may be inconsistent with autonomous choice (for instance, those consistent with false or oppressive norms).

Of the two prominent substantive views considered, one (Stoljar's) posited the value of the true and the good or non-oppressive; another (from Hill and Baron) posited the value of self-respect as necessary for autonomy. However, though motivated to make discriminations about certain intuitively problematic choices in oppressive social contexts, I argued that neither view succeeded in capturing the stated intuitions. Stoljar's condition was explanatorily inadequate, whilst neither Hill's nor Baron's self-respect requirement could diagnose the problem with the deferential individual.

In this chapter, I give a more general argument, to the effect that no account that incorporates a substantive condition of the following form should be accepted:

(RelChoice) A necessary condition for autonomous choice is that the agent's choice is in accordance with value v .

I will argue that such a condition cannot make the kind of intuitive discriminations between autonomous and non-autonomous agents required: it faces 'the problem of false negatives', namely, that seemingly autonomous choices are diagnosed as non-autonomous. Thus the motivation to accept such a relational condition falls away. I argue for this conclusion by noting that the condition as it stands is underspecified. Once fully spelt out, the problems that face such a condition become apparent.

We will see the concern, from Christman, that conceptions of autonomy that incorporate value are not able to play the normative roles (of grounding respect, protecting from paternalism, and so on). Having considered a plausible line of response to this concern, I set out my argument for the claim that substantive accounts cannot capture the intuitions that proponents of such views set out to make sense of. This is important, because the substantive theorist might respond to the first objection that she is concerned with a different notion of autonomy. Irrespective of the legitimacy of this move, it will be of no use if the substantive condition can't capture the intuitions that motivate theorists to posit it.

4.1 A concern about values

John Christman has argued that substantive accounts should be rejected, because the value-laden content that is incorporated into the conditions for autonomy means that they are unable to play the roles in the normative frameworks that account in part for its value. These roles are part of the framework – consideration (A) – set out in Chapter 1 that enables us to evaluate a conception of autonomy. His concern is that such a value-laden conception brings with it the danger of “exclud[ing] from participation those individuals who reject those types of social relations demanded by those [substantive] views”.¹ With regards to the claim that an agent who fails to choose in accordance with the specified value is non-autonomous, Christman remarks:

To say that she is not autonomous implies that she does not enjoy the status marker of an independent citizen whose perspective and value orientation get a hearing in the democratic processes that constitute legitimate social policy.²

Suppose an agent meets all the content-neutral conditions but fails to choose in accordance with the substantive value specified. Then, Christman claims:

despite her authentic, competent, and ‘sober’ [choice] (by hypothesis), her lack of autonomy ... would allow other agents and representatives of coercive social situations to intervene to relieve her of this burden and to restore her autonomy (at least in principle). This implication should be troubling.³

¹ Christman, J. (2004) *Op.Cit.* p.156.

² *Ibid.* p.157.

³ *Ibid.* p.157.

Christman's concern is that it is troubling to hold that, simply because an agent does not hold a certain value commitment, she cannot be a participant in political processes, and is not protected from paternalistic interventions. Now, Christman's worry about value-laden accounts is one that he thinks faces all constitutively relational conceptions, and I shall return to address this aspect of his worry in Chapter 7.⁴ For present purposes, however, the central worry is that insofar as a conception of autonomy is value-laden, it cannot stand in the normative frameworks required of it. Individuals are due respect not because of what they value, but because they have, for example, the capacity to value, and to deliberate about what to value. They are candidates for political participation not because of certain value commitments, but because of their ability to deliberate about what conception of the good to pursue. This thought underpins the claim that such agents should not have their actions unduly interfered with,⁵ even when the interference is motivated by concerns for their own good. It would be dangerous, and antithetical to an appreciation of a pluralistic set of values,⁶ to delineate the class of agents who have these normative benefits according to whether they choose in line with specific values. This aspect of Christman's concern seems right, and insofar as the substantive theorist cannot respond to it, this seems to provide good reason to reject such views. It is worth briefly considering some of the responses to this concern, however.

4.1.1 A different stripe of autonomy?

The substantive theorist might well accept Christman's claim that a value-laden conception of autonomy should not do the work of establishing entitlement to normative benefits. She might claim that the term 'autonomy' is broad, and can mean different things, and that Christman's notion of autonomy differs from that which is constrained by substantive values.⁷

⁴ To anticipate: I shall argue that he is wrong to claim that all constitutively relational conceptions are value-laden.

⁵ Of course, there are some exceptions: intervention to prevent harm befalling others may be justified.

⁶ This might be understood as a 'reasonable pluralism' in Rawlsian terms. The fact of reasonable pluralism is stated as the fact that there are an incompatible set of comprehensive views (accepted values, religions, aims and goals) each of which is the result of good reasoning about what comprehensive view to adopt and how to shape one's life. See Rawls, J. (1993) *Political Liberalism* Columbia University Press, Esp. Lecture 1, pp.4-46.

⁷ See Arpaly, N. ((2003) 'Varieties of Autonomy' in her *Unprincipled Virtue*, Oxford University Press, pp.117-149) for discussion of some of the different senses of autonomy present in the literature.

One form this response can take is to claim that substantive values are relevant to autonomy in a ‘personal authenticity’ sense. Recall that in Chapter 1 we saw the distinction between what Scanlon called ‘the political problem of free will’ – namely, the problem of the extent to which agents’ choices and conceptions of the good should be respected, protected from paternalistic intervention, not coerced by the state, and so on – and ‘the personal problem of free will’. This latter problem is, in part, a matter of how to make sense of the causes that leave agents feeling ‘alienated’ from their actions – feeling “manipulated, trapped reduced to the status of a puppet”.⁸ The proponent of a substantive condition may claim to be concerned with a similar, personal authenticity sort of autonomy; with how socially oppressive contexts might leave agents feeling alienated from their actions in such a way. A substantive condition, she might claim, is well placed to diagnose cases in which there is, intuitively, something problematic with an agent’s autonomy, or ‘personal authenticity’; namely, when she is choosing in a way that does not accord with the specified value (perhaps, the theorist will go on to claim, this is a value to which all agents are, authentically, committed). It does not follow from diagnosing a lack of autonomy in the ‘authenticity’ sense that we can intervene with the agent’s choice, and so on. But it does follow that we can make some critical comments about the social contexts that thwart or make difficult this personal authenticity.

Or, a theorist may claim, as Oshana does, that the objector is concerned with a ‘thinner’ sense of autonomy than that which incorporates value-constraints, which is ‘personal autonomy’ properly understood. She writes:

it is not personal autonomy that is the subject of the liberal’s concern as Christman presents it, but autonomy of some other (political) variety. ... But accounts of political autonomy are generally too thin to illuminate a full account of self-determination. Personal autonomy is a marker of one of a wider range of lifestyles political liberalism accommodates, but there may be politically autonomous agents who fail to be personally autonomous in the manner I describe.⁹

On this view, the concern about value-ladenness only holds for those accounts intended to characterise a ‘thin’ notion of autonomy – one that characterises the markers that citizens must manifest to be candidates for participation, protected from interference,

⁸ Scanlon, T.M. (1986) *Op.Cit.* p.157. As I have mentioned, he is concerned in particular with the way that we can understand some, but not all, of our actions in this way, given the causal thesis.

⁹ Oshana, M. (2006) *Op Cit.* p.102. Note, however, that Oshana at other points insists that she is concerned with the same notion of autonomy (see her discussion of the ‘apples and oranges’ objection, (at pp.93-93) to which I refer in the next chapter).

and so on. But insofar as one is seeking to give an account of a richer phenomenon – “a full account of self-determination” – it may yet be legitimate to incorporate substantive values. Oshana writes that “it is one thing to say there are a variety of social arrangements that comport with well-being [to wit: those that individuals should be protected from undue intervention in so choosing], quite another to claim that each of these arrangements will provide a life of autonomy”.¹⁰ On this view, it does not follow from diagnosing a lack of autonomy in the ‘full self-determination’ sense (failure to choose consistently with the substantive value) that the agent fails also to possess the normative benefits to which she is entitled by her possession of the thinner stripe of autonomy.

In these ways, then, the relational theorist might respond to Christman’s concerns about value-ladenness. This line of response faces challenges, not least in explaining the value and importance of the notion of autonomy – in the personal authenticity, or ‘full’ self-determination sense – once it has been hived off from the normative benefits that accrue to the notion of autonomy the objector is concerned with. I return to address some of the issues raised by Oshana in Chapter 6. Here, however, I show that neither of these lines of response will serve to help the substantive theorist. In the rest of this chapter, I block lines of response such as the two above, by showing that the substantive condition cannot make good sense of the intuitively non-autonomous choices. She cannot adequately discriminate between autonomous and non-autonomous choices, as is the stated aim of such views, whatever kind of autonomy or authenticity she purports to be concerned with.

4.2 Three ways to understand the substantive condition

So far, we have seen that substantive conceptions of autonomy supplement the necessary content neutral conditions with a condition of the form:

(RelChoice) A necessary condition for autonomous choice is that the agent’s choice is in accordance with value *v*.

The thought is that whatever content-neutral conditions are deemed to be the correct ones, it is only in conjunction with a substantive condition of this form that we have

¹⁰ *Ibid*, p.104.

conditions sufficient for autonomy.¹¹ As stated, however, it is not clear precisely what the substantive condition demands: there are, as I will spell out, three ways of understanding it.

Consider, first, Christman's characterisation of substantive conditions, according to which they:

demand that particular values or commitments must be part of the autonomous agent's value or belief corpus.¹²

This claim suggests that the agent must believe that *v* – the value specified by the account – is a value (worth promoting/preserving, reason-giving, and so on). And, presumably, this belief must constrain the choice of the agent. If it is the epistemic states of the agent that are of interest, then one natural way to interpret (RelChoice) is:

(ERC) A necessary condition for autonomous choice is that the agent is committed to value *v*, and believes that her choice is in accordance with value *v*.

Compare this rendering of the condition with Stoljar's claim, in her discussion of the alleged non-autonomy of women who chose in accordance with 'false and oppressive' norms. There, she claims that

because of the internalisation of the [false] norm, they do not have the capacity to perceive it as false.¹³

I have earlier argued that it is very difficult to make any sense of this claim. However, reading between the lines, we can see that what Stoljar has in mind here as constraining choice surely cannot be (ERC). Why not? Because Stoljar claims that the agents who accept the 'false norm' cannot see its falsity. Accordingly, some agents who act upon 'false norms' will take themselves to be acting in accordance with the true and the good – it is plausible that such agents will believe themselves to be acting in accordance with the value specified by Stoljar's substantive account. Perhaps what matters, with respect to the autonomy of the agent's choice, is whether or not they are in fact choosing and acting in accordance with the true and the good.

¹¹ See Chapter 2 for a spelling out of some of the content-neutral conditions to which substantive conditions are added.

¹² Christman, J. (2004) *Op.Cit.* p.148.

¹³ Stoljar, N. (2000) *Op.Cit.* p.108.

This suggests there is a second way of reading the condition, according to which the agent's choice must be 'world-guided', in that it must in fact accord with the value specified by the account:

(WRC) A necessary condition for autonomous choice is that the agent in fact chooses in accordance with value v.¹⁴

Substantive theorists are not clear about which of these understandings of (RelChoice) should be adopted. Some accounts appear to hold a combination of the two views. In the diagnosis of what is wrong with the deferential individual, for instance, Hill claims that

the objectionable feature ... is his tendency to disavow his own moral rights [i.e. the right not to defer] either because he misunderstands them or because he cares little for them.¹⁵

The claim here concerns both what the agent in fact chooses (the disavowal of her rights), and how he understands the choice (in light of his false beliefs). Hill seems to be demanding that a) an agent commit to value v (self-respect, derivative from the moral law, in this case); b) believes herself to be acting in accordance with this value; and c) that this belief be true; the choice is in accordance with the value. Part of choosing and acting in accordance with this value is having true beliefs about whether the choices in question so accord: having false beliefs demonstrates a failure to properly appreciate the value.¹⁶ Thus on the table is also a hybrid understanding of the substantive condition:

(HRC) A necessary condition for autonomous choice is that the agent commits to the value v, believes her choice to be in accordance with value v, and in fact is in accordance with this value.

¹⁴ What is it to choose in accordance with some value? Again, this is often left unspecified. For present purposes, I take it that to choose in accordance with value v has two readings: weak reading – the choice does not conflict with (undermine, thwart) the promotion of this value; strong reading – the choice promotes the value (i.e. brings about more of it, or more appreciators of it). I am primarily concerned with the weak reading here, although my claims will apply to the stronger understanding also.

¹⁵ Hill, T. (1987a) *Op. Cit* p.13.

¹⁶ It seems that there will be certain kinds of false belief that are not inconsistent with properly understanding the value. Consider a choice for slavery. Hill envisages a case where the agent misunderstands the value of her own humanity, and hence the value of self-respect, and so believes slavery to be compatible with this value. But there is another possibility: the agent might simply misunderstand what slavery involves. It is not clear, in cases such as the latter, that the agent also misunderstands the value of self-respect. It seems that she understands this value properly, insofar as she meets the following counterfactual condition: were she to acquire true beliefs about the nature of slavery, she would recognise that it is incompatible with the value of self-respect (and subsequently change her choice so that it accords with this value).

Having separated out these dimensions of the substantive accounts, in the following sections, I explore the prospects for each understanding.

4.3 Epistemic conditions for autonomous choice

The first way of understanding the condition is:

(ERC) A necessary condition for autonomous choice is that the agent is committed to value v , and believes that her choice is in accordance with value v .

This condition places an epistemic constraint on autonomous choice: the agent must have a certain belief about the relation between her choice and a certain value. That is, the agent must have, in her belief corpus, the following:

(B) My choice is in accord with value v / is not in conflict with value v .

Note that the agent must also be committed to the value. This commitment might involve caring in a certain way about the value, and also holding a belief about the value, such as:

(C) V is valuable: worth pursuing, reason-giving, to be promoted, preserved etc.

Adding the demand that the agent believes (C) also means not only that the agent must act in accord with the value, but also in fact be committed to the value. It appears that a substantive account should demand (C) as well as (B). Suppose that the good is the value in question, and one aspect of this is health. An agent chooses to eat papaya, believing this to be in accordance with the value of health. But her choice is not motivated in the slightest by health considerations; she does not take health to be something worth pursuing, and is merely concerned to satisfy her palate. In such a case, the fact that her choice is in accordance with the value may play no role in her coming to make that choice. (I return to such considerations below.) Thus there appear to be two epistemic demands on autonomous choice: the agent must believe both (B) and (C). How plausible is it to accept these epistemic conditions? By way of answer, I first take a

detour into the literature that discusses the plausibility of epistemic conditions for autonomy.

4.3.1 Other epistemic conditions

McKenna claims that there are no epistemic conditions on autonomous choice.¹⁷ If he is right the substantive theorist faces serious problems in demanding that agents, in order to be autonomous in choosing, believe (B) and (C), above. McKenna takes himself to be disagreeing with Mele, who argues that there is an epistemic condition for autonomy. In the following I consider what we can learn from their arguments for and against epistemic conditions for autonomous choice.

a. McKenna vs. Mele on epistemic conditions for autonomy of choice

McKenna holds that there are no epistemic conditions for autonomous choice, at the stage of deliberations and subsequent making of choice.¹⁸ The example he uses to support this claim is as follows (I paraphrase):

Tal recognises that the unconscious Daphne needs the medicine called The Good Stuff. He gives her a dose from the bottle marked 'The Good Stuff'. Due to a pharmaceutical hitch that is nothing to do with Tal, this bottle in fact contains The Bad Stuff, and Daphne dies.¹⁹

McKenna's claim is that Tal's autonomy of choice is not compromised by his false belief (although we would excuse him from moral responsibility). There is nothing wrong with the way that Tal's deliberations are conducted, and he governs his deliberations, judgements, subsequent choice and action in a way consistent with good agency. Thus he is autonomous in his choice of action – to assist Daphne by giving her medicine – despite his woefully false belief (in virtue of which it is inappropriate to hold him morally responsible for Daphne's death).

¹⁷ McKenna, M. (2005) 'The Relationship between Autonomous and Morally Responsible Agency', in *Personal Autonomy*, Taylor, J.D. (ed.), Cambridge, Cambridge University Press, pp.205-235.

¹⁸ McKenna does accept that there might be some epistemic conditions on the formulation of the policies and principles that one utilises in making choices. If one is seriously misinformed when formulating the policies that one later utilises, this might undermine autonomous choice. But the constraint is not on the beliefs of the agent *at the time of choice*, but on the beliefs of the agent *at the time of formulating principles*. (One might worry that it is somewhat implausible to think of these processes as clearly distinct, but I set aside these concerns for now).

¹⁹ See McKenna, (2005), *Op.Cit.* pp.208-211.

McKenna takes himself to be disagreeing with Mele, who, by contrast, claims that there is an epistemic condition on autonomous choice and subsequent action. He claims that an agent must be in an ‘informational state’ such that she can “make informed decisions about how to pursue [her] ends” (of course, being in a position to make informed decisions does not require that the agent have no false beliefs; just that a significant portion of her relevant beliefs are true).²⁰ An agent should be in an epistemic state – have sufficiently many true beliefs – such that she has control over the success of achieving her ends. Absolute control would be too much to demand; rather an agent should have control such that her choice of action “increases the likelihood of [achieving her ends] above that provided by mere chance”.²¹ In cases in which an agent is so misinformed that executing her choice will in fact thwart her aims, the misinformation means lack of autonomy. Consider Tal: he has false beliefs such that his executing his choice decreases the likelihood of him achieving his aim of saving Daphne. This false belief renders him unable to control the relative success of his actions, and as such renders him non-autonomous, according to Mele.

b. Resolving the disagreement

Here we have two views; according to one, there is an epistemic condition for autonomous choice; according to the other there is not. The first thing that we should note, however, is that the disagreement is not straightforward. Mele and McKenna are in fact talking about different things, but using the term ‘autonomy’ to refer to both. McKenna is concerned with whether or not an agent is playing a governing role with respect to her deliberation, and the choices made on the basis of this deliberation. Mele, on the other hand, is concerned with whether the agent is able to be effective or successful in executing and achieving her aims through her action. Thus it appears that he is concerned rather with the agent’s personal efficacy in action.

The bearing of this on their dispute is that there need not in fact be any disagreement: it is possible and indeed plausible to claim that there is no epistemic condition on governance of deliberation and choice, but that there is an epistemic condition on personal efficacy of action – an agent will be thwarted in achieving her goals if she has certain false beliefs. An agent with some false beliefs may be a perfectly functioning deliberator and chooser, given the deliberative input she has – it’s just that

²⁰ Mele, A. (1995) *Autonomous Agents: from Self-Control to Autonomy*. Oxford, Oxford University Press. See esp. pp.177-182.

²¹ *Ibid*, p.181.

she won't be very effective in successfully achieving her ends when acting upon her choice if the choice is informed by false beliefs. It is plausible that we could accept both claims, then: that there is no epistemic condition for autonomous choice and that there is an epistemic condition on the efficacy of autonomous action.

c. A further disagreement

It is worth noting that at some points Mele does appear to be concerned with the deliberations and choices of the agent. In the context of examining the kind of 'informational state' that autonomous agents possess, Mele asks:

is it plausible that although [the systematically misinformed agent] is not autonomous in his efforts to [achieve his ends], he nevertheless is autonomous with respect to his deliberation about the means of doing this? Not at all. Deliberation is informed by the deliberator's beliefs. By controlling what [an agent] believes, [one] gains control over what deliberative conclusions [the agent] reaches.²²

Here Mele is concerned with whether an agent can be autonomous in deliberation and choice whilst possessing false beliefs – and we can presume the false beliefs in question to be relevant and wide ranging, given the systematic misinformation of the agent in the example. In focusing on the agent's autonomy in her deliberation and choice, Mele is here talking about the same aspect of autonomy as McKenna. So they are disagreeing, insofar as Mele holds that autonomous agents do meet certain epistemic conditions. On what basis does Mele argue for an epistemic condition for autonomy in deliberation and choice? In being misinformed, Mele claims, an agent lacks control over her beliefs; others have control over the content of her beliefs, and hence her deliberation and subsequent choice. Such an agent is non-autonomous.

Should we accept Mele's claims on this point rather than McKenna's contention that there are no epistemic conditions on autonomous choice? The considerations I raise here will not be decisive, but I suggest that we should reject the case that Mele makes for an epistemic condition. There are two points here. The first pertains to what is doing the work in Mele's example: whether it is the deception by others, or rather the agent's state of having false beliefs. If it is the deception by others that is driving Mele's claim that the agent is non-autonomous, then his is an argument for the incompatibility of manipulation and autonomous choice, rather than for an epistemic condition for autonomy.

²² *Ibid*, p.182.

The claim that being systematically misled undermines autonomy can be held whilst denying that there are epistemic conditions for autonomous deliberation and choice; it is the ‘being misled’ that undermines autonomy, not the false beliefs that (may) result.²³ Whilst it is intuitive to say of an agent who has been systematically misled that ‘her autonomy is undermined’ or that ‘she has had her autonomy violated’, the sense of autonomy in these utterances is quite different to that of governance of deliberation and choice.²⁴ That is, to say that an agent’s autonomy is violated in such a context is not to say that her capacities for deliberation and choice are tampered with. Rather, it is to say that she has not been treated as she ought to have been.²⁵

However, we should not understand Mele in this way, because he accepts that natural phenomena (such as the electromagnetic fields of a Bermuda Triangle) may cause the agent to be in a state of misinformation in a way that undermines autonomy. This leads us towards an understanding of his claims as concerning how the possession of false beliefs might in itself undermine autonomy of deliberation and choice.

Let us suppose, then, that it is not the misleading that is doing the work in the above example, and rather the false beliefs that the agent has. This brings us to the second point: thus understood, should we accept Mele’s claims, above, as persuasively showing that agents with relevant false beliefs lack autonomy in deliberation and choice? The answer to this should surely be no. The consideration Mele appeals to here is that others are controlling the input (beliefs) to the agent’s deliberation.²⁶ But no agents have direct control over the input to their deliberation. Insofar as agents do not

²³ Note that an incompetence deceiver who has false beliefs may ‘mislead’ an agent into believing truths.

²⁴ This point is made by Hill, in ‘The Importance of Autonomy’ (1987c). His claim is that such talk of autonomy violation is best understood by conceiving of this sense of autonomy as a right; ‘it is to *grant* the person a right to control certain matters for himself or herself’ (p.48). When deceived or manipulated, we can say that whilst the agent’s autonomy qua legislator, or qua competent agent, is intact, her right to control her belief formation has been violated. Even this, however, cannot be quite right: as I claim below, we should not demand that agents have (direct) control over the input to their beliefs, or the belief-formation process.

²⁵ See Buss, S. ((2005) ‘Valuing Autonomy and Respecting Persons: Manipulation, Seduction and the Basis of Moral Constraint’, *Ethics*, 115, pp.195-235) for claims to this effect, and how this challenges normative systems that claim that autonomy is the value that generates obligations. My claim does not address this latter claim; I merely emphasise that claims about autonomy violation are not, in this instance, claims about the violation of an agent’s capacities for rational deliberation and choice.

²⁶ The nature of the control is relevant: if an agent is very regularly ‘fed’ with false beliefs, this appears more autonomy undermining than if she is given false beliefs and then left alone to deliberate on the basis of them. But these intuitions do not support the claim that false beliefs undermine autonomy, but again, that manipulation does.

have direct control over the forming of beliefs, they have little control over the input into deliberative processes. Moreover, agents should not have this kind of control. On occasions on which agents do exert control over their beliefs – engaging in wishful thinking, self-deception, confirmation-bias, and so on – we take them to be guilty of epistemic vice.²⁷

Thus Mele's argument for an epistemic condition seems problematic in the following two ways: first, he relies upon an equivocal use of 'autonomy' in order to garner intuitive support for the claim that systematic deception undermines autonomy of choice. Second, his discussion appears to suppose that lacking control over the input to deliberation undermines autonomy in deliberation. But this is implausibly strong: even ideally self-governing deliberators could not achieve this and, moreover, they should not. This argument from Mele cannot establish that there is an epistemic condition for autonomous choice that demands the agent's (relevant) beliefs be under her control.

4.3.2 Back to the substantive condition

We have seen, then, that it might be plausible to hold that there are epistemic conditions on being efficacious in action such that one is effective in achieving one's ends. But it is more difficult to sustain the claim that agents must not have false beliefs if they are to be autonomous – that is, self-governing – in their deliberations and choice. Indeed, we have intuitive support for the absence of such epistemic conditions, from McKenna's examples, and no good argument for such conditions from Mele. How might these preliminary conclusions be relevant to the assessment of the epistemic condition (ERC) in the substantive account of autonomous choice?

²⁷ Further, it is particularly odd that Mele should take this line about control over deliberative input, as elsewhere (in his argument for agnostic autonomism (in (2005) 'Agnostic Autonomism Revisited' Taylor (ed.) *Personal Autonomy*, pp.109-124)) Mele *relies* upon the claim that agents lack control over this input in his argument for a plausible libertarianism. His libertarianism is plausible, he claims, because it can withstand the objection that indeterminism in the agent's mental states or actions means that she lacks control (it is random, or chance, which mental state she has, or choice she makes, or action she performs). Mele's proposal is that, because the beliefs that we start deliberation from are mental states over which we have no control (often they just 'pop into' our heads), if it turns out that there is indeterminism at this stage, then the libertarian account that commits to indeterminism places in the agent's hands no less control than that which the determinist accords her (whilst what thought pops into the agent's head is determined, it remains something that is not under her control). Thus the libertarian can avoid the concern that she gains freedom from determinism at the cost of relinquishing control.

If McKenna is right that there are no epistemic conditions for autonomous choice, then this seems to be bad news for this rendering of the substantive condition. But the first thing to notice is how different the conditions under consideration by Mele and McKenna are from the kind of condition we have with (ERC). Mele and McKenna are concerned not with whether autonomy requires any specific belief, but with whether a certain portion of the agent's beliefs must be true if the agent is to govern her deliberations and choice. The epistemic condition specified by the substantive account would be quite different, requiring that the agent believe a very specific propositional content, of the form:

(B) My choice is in accord with value v.

The views we have considered then, such as Mele's, could lend little support for a view that posits this kind of epistemic condition alone. Rather, if there are epistemic conditions for autonomous choice – a big if, given that Mele's argument for this claim was a bad one – this will only support the substantive theorist's claim that agents must believe (B) in cases in which that belief is true. The substantive theorist gains no support for the belief simpliciter, but rather support for the belief's role qua true belief.

Even then, given that not all of our true beliefs are relevant to our choices, we would have to see why this specific belief is relevant. Many of my true beliefs – for example, my belief that clown loaches are scaleless fish – are totally irrelevant to both my deliberation and choice on many occasions. So an explanation of why the particular belief about coherence with some value is relevant needs to be given. At this stage, then, all we can claim is that the burden of proof to show that specific beliefs of this form are relevant to and necessary for autonomous choice resides with the substantive theorist. The two prominent accounts that I assessed in the previous two chapters failed to meet this burden.

It is worth noting, moreover, that the implications of taking the (ERC) condition plus the content-neutral condition to be sufficient for autonomous choice are highly counterintuitive. In cases in which the agent's choice is not in accordance with the value specified by the account, the epistemic condition will demand that the agent form or sustain false beliefs, in order to be autonomous. The arguments for epistemic conditions that we have looked at are all concerned with whether the agent must have true beliefs. Considerations there were inconclusive. It seems even more problematic that a

substantive account might require that agents hold false beliefs. It is one thing to hold that false beliefs can be compatible with autonomous choice. It is quite another to claim that the addition of a false belief to a motivational set that meets all the content-neutral conditions is enough to make autonomous a choice that stems from such a psychology. Whilst it is not clear that true beliefs are necessary for autonomous choice, it is surely problematic to hold that content-neutral conditions plus a false belief is sufficient. There is no support for this kind of claim from the considerations raised by Mele.

Let us see what would be demanded by an epistemic rendering of the substantive condition, if we plug in the values specified by the accounts we have looked at. A substantive account may demand that the agent believes:

(KB) My choice is in accord with value of self-respect/ is not in conflict with value of self-respect

Suppose that an agent chooses slavery. An account that took the form of (ERC) would demand that, as a necessary condition for autonomy, the agent believe that her choice is in accordance with the value of self-respect. Let us for now grant the claim of the Kantian substantive theorist that the choice for slavery is not consistent with the value of self-respect. Then the requirement that the agent believes her choice to be consistent with the value of self-respect would be a demand for the agent to form or sustain a false belief. This is highly counter-intuitive: even if true beliefs are not required, a requirement for false beliefs surely cannot be what makes the difference to an agent's choice being autonomous or not.

I have noted that Stoljar's substantive condition should not be understood as an epistemic condition. Indeed, we can now see just how problematic it would be for an account that specifies as the value that the agent must believe herself to be choosing in accord with 'the true and the good' to take the form of (ERC). For such an account, meeting the epistemic condition might sometimes require falsely believing that one's choice is in accordance with the true and the good. An account of self-governance that gives central place to the value of the true and the good in autonomous choice is self-defeating if it requires agents to sustain false beliefs in order to be autonomous with respect to her choice.

Again, whilst holding false beliefs may not be incompatible with autonomous choice, it is surely theoretically problematic to hold that a false belief is what pushes a

deliberatively competent agent over the threshold for autonomy. Such a demand is so counter-intuitive that we can conclude this necessary condition – jointly sufficient with the content-neutral conditions – cannot be getting at what is taken to be important by substantive theorists. Surely what is important is that the agent's choice is in accordance with this value. This directs us to consideration of the world-guided understanding of (RelChoice), to which I turn in the next section.

4.4 World-guided conditions for autonomous choice

The world-guided understanding of the substantive constraint on autonomous choice was set out as:

(WRC) A necessary condition for autonomous choice is that the agent in fact chooses in accordance with value v.

This understanding of the substantive account will not be focused on what the agent believes to be the case with respect to the relation between her choice and the value v. Rather, it will be focused on what the relation in fact is between the agent's choice and this value. Accordingly, the demands that the epistemic rendering of the condition seemed to be committed to, in terms of requiring false beliefs to be necessary for autonomy, can be avoided.

On this understanding, what is required is that the choice does accord with the value. Accordingly, a consequence of this kind of substantive condition is that the beliefs of the agent are not relevant to an assessment of her autonomy. Rather, what matters is simply whether or not the choice in fact is in accordance with the value specified by the account. However, the constraint specified by (WRC), in conjunction with whatever content-neutral conditions are demanded for autonomous choice, cannot be what the substantive theorist has in mind as jointly sufficient for autonomy. To see why, consider the following case:

Mistake: Claude believes that not dropping litter accords with the value of the good, in particular, the value of maintaining a clean and pleasant environment. But she doesn't care about promoting the good – in particular, she doesn't care about keeping her locality clean and pleasant, so she drops her banana skin in the park – an action which she believes conflicts with the value of the good. But she is mistaken: in dropping the fruit skin, she in fact contributes a good amount of phosphorous and potassium to the soil, thus improving the quality of planting soil in her local park. In choosing to drop

litter, then, on this occasion Claude chooses and acts in accordance with the value of the good.

In Claude's case, the relation between her choice and the value *v* is purely accidental. It is a matter of fluke that she chooses in accordance with the value. Now we are unlikely to judge an agent to be non-autonomous. But insofar as the choices in accordance with some value are important to the substantive theorist, this is surely not the role that is envisaged for the value. This shows us, first, that the world-guided condition ought to be modified to incorporate also the demand that the agent commit to the value in question:

(WRC*) A necessary condition for autonomous choice is that the agent is committed to the value *v*, and in fact chooses in accordance with value *v*.

Second, mistake shows us that an epistemic condition appears, after all, to be required. The world guided condition, with whatever content neutral conditions are necessary, are not jointly sufficient. Third, this kind of case suggests that an adequate substantive condition must say something about the relation of the agent's belief to the choice that she makes. The substantive theorist needs to demand that the agent who chooses in accordance with the value, *v*, does so because of her (true) beliefs about the coincidence of her choice and the value (rather than due to fluke).

In Claude's case, the problem is that she has false beliefs but her choice nonetheless in fact accords with the value. But the problem runs the other way also, and in a way that poses more problems for the substantive theorist. Consider the following case which challenges the necessity of the world-guided aspect of the condition, in which the agent's choice fails to accord with the value in question:

Misfortune: Suppose for present purposes that the value, *v*, is specified as self-respect. Gus, a Kantian, believes that only an action that is morally right is consistent with self-respect. He faces a moral dilemma, and no matter how long he deliberates, he cannot determine what his moral duty consists in – it seems that either option involves a violation of humanity. He takes option A, an option which is not in accordance with self-respect, but which, after careful deliberation, he has decided is the way to act, in the difficult circumstances.

Despite his careful deliberation, if (WRC) is a necessary condition, the substantive theorist must hold that Gus does not act autonomously, because his choice in fact fails to accord with the value of self-respect. This is counter-intuitive: the unfortunate circumstance may mean that Gus is not morally blameworthy for so acting as he might

otherwise be. But the circumstance does not make his choice non-autonomous; indeed, such a challenging case, that demands careful deliberation, makes us more inclined to regard his choice as autonomous; it certainly manifests a higher degree of deliberation than many run-of-the-mill choices.²⁸ Note that, were Gus in a different scenario, in which the options were very slightly different, such that (unbeknownst to Gus) one option was morally required, his choice (if it coincides with this option) would be autonomous. That a change in the world might make all the difference to Gus' autonomy seems wrong. Next, consider:

Rude: Suppose for present purposes that the value, v , is specified as 'the true and the good': an agent's choice must in fact accord with the value of 'the true and the good'. Suppose that Betty cares a great deal about observing etiquette, believing it to be part of the good to follow the norms of polite behaviour. But she has false beliefs about what, in a particular context, she ought to do (suppose that the etiquette book she consulted was out of date). She believes that it is polite to belch loudly after eating. And suppose that this is not the case: so she has false beliefs about the norms of etiquette in such context. When, after the meal, Betty belches, she believes that belching is in accordance with the true and the good. But her choice to do so is not in fact in accordance with the true and the good.

Betty believes that her action is in accordance with the true and the good, but it is not. Nonetheless, whilst we might excuse her on the grounds of having made a mistake, we would not hold that she was non-autonomous in so acting. Despite choosing and acting in a way that did not accord with the true and the good, she maintained self-governance of, or control that she has over, her deliberations and subsequent choice. But, insofar as the (WRC) condition is necessary and is not met, Betty is not autonomous in her choice. Note, however, that had Betty been in a different context – one in which the norm is operative, and it is in fact polite to belch after eating – she would have been autonomous in her deliberation and choice. The difference in the world – in the operative norms in a particular context – are what makes a difference to her autonomy, in such a case. This again, is counter-intuitive.

Insofar as the fact of the matter regarding the coincidence of the choice with the value is a condition for autonomous choice, autonomy seems hostage to the world in an undesirable way. When agents are in good shape, with respect to beliefs formed on the basis of evidence, and with respect to deliberation, it is counter-intuitive to suppose that

²⁸ Note though that he may fail to be autonomous in the Kantian sense of full moral autonomy, insofar as he does not act in accordance with the moral law. But moral autonomy is not the sense of autonomy at issue here.

unfortunate or unexpected features of the world could undermine their autonomy of choice. Of course, too many of these false beliefs, and the agent's personal efficacy in achieving her ends will be thwarted. But as I have already addressed above, this is a different matter from that of autonomous choice.

To hold autonomy of choice hostage to the world in this way is problematic. Thus it is implausible to demand that the world-guided condition is necessary for autonomy in choice. We will see that this is ultimately deeply problematic for the substantive theorist: in order to make the diagnoses required, the world-guided condition has to do all the work in a substantive account. Given the implausibility of the (WRC) condition, this is a serious problem for the proponent of the substantive account.

To recap: Mistake indicates that the world-guided substantive condition, in conjunction with the content neutral conditions, cannot be sufficient for autonomy. Whilst the epistemic condition (ERC) was problematic in demanding that false beliefs, on occasion, be sustained in order to achieve autonomy in choice, the considerations raised here show that the agent's epistemic states are in some way relevant to the autonomy of the agent in choosing. The agent's beliefs about the coincidence of their choice with the value must be relevant. Moreover, the considerations raised by Misfortune and Rude put pressure on the world-guided condition being necessary for autonomous choice.

This pushes us towards the thought that the substantive theorist is after a hybrid condition that incorporates both epistemic and world-guided conditions, and speaks to the connection between the agent's beliefs about the relation of their choice to the value in question, in a way that can avoid the problems raised in this and the previous sections. In the following, I set out what this condition will have to look like, and then outline problems that this view faces, given what motivates substantive conditions.

4.5 A hybrid condition

A purely epistemic reading of the substantive condition leads to counterintuitive conclusions with respect to autonomous agents sustaining false beliefs. This surely cannot be a requirement on effective self-governed choice. And one of the problems with a purely world-guided rendering of the substantive condition was that it is unable to capture the kind of connection that is surely required between the agent's choice and

their beliefs about the coincidence of that choice with some value (as specified by the substantive account). Perhaps such problems can be avoided by incorporating both epistemic and world-guided considerations.

Here is a first attempt at a more adequate specification of the substantive condition:

- (HRC): a necessary condition for autonomous choice is that
- a) the agent is committed to value v ,
 - b) the agent believes that her choice is in accordance with this value, and
 - c) her choice in fact is in accordance with this value

The considerations raised by mistake indicate that we should also add:

- d) The agent's beliefs about the value must play a causal role in the agent's choice.

This condition seems to do a better job of capturing what the substantive theorist seeks. It captures the thought that, if what is important is that the agent chooses in accordance with the value specified, the agent should also be committed to this value, and there should be a connection between the agent's belief about the prospective choice and the value, and her making of this choice. No substantive account has drawn out the precise nature of these requirements, when offering claims of the form:

(RelChoice) A necessary condition for autonomous choice is that the agent's choice is in accordance with value v .

I have set out in considerable detail what the proponent of a substantive account appears to be committed to as true of autonomous agents, with respect to their commitments, beliefs, and the role of these states in causal explanations of their choice. These more detailed demands, it seems, must be incorporated into a substantive account – otherwise the agent need not have proper appreciation for the value specified, nor need their beliefs or appreciation of the value play any role in the agent's choice.

Having set out what substantive accounts must be committed to, if they are to have a plausible rendering of the way that the specified value enters into the conditions for autonomous choice, in the next section I argue that given these commitments, substantive accounts cannot play the role that has been intended, in terms of making sense of what is intuitively problematic about certain agent's choices.

4.6 Problems with the fully specified substantive condition

Recall that the stated aim of the substantive theory of autonomy is to make sense of intuitions about which choices are non-autonomous. Specifically, the substantive condition is posited by way of making sense of what is particularly problematic about certain choices (for contraceptive risk, for deferential roles, say). Note, then, that the substantive theorist is not merely after a diagnosis (of the agent's choice as non-autonomous): she is after a distinction – between these troubling choices, and those that are intuitively autonomous. Now, this means that there is a presumption that a range of choices are autonomous, in a way that the particular choices at issue are not.

Now, the substantive theorist is motivated to capture the following two intuitions, each of which has informed the accounts that we have considered in previous chapters, respectively:

(I1) choices that are informed by and accord with oppressive norms are non-autonomous.

(I2) choices for deferential roles are non-autonomous.

The key thought is that, distinctions should be drawn between agents whose choices are relatively unproblematic, and those agents who choose in accordance with oppressive norms. For example, a substantive account should be able to distinguish between the cases supposed to be intuitively problematic:

(1a) Lana chooses a role of deferential housewife

(1b) Santa chooses to take contraceptive risk, because she believes that it is unfeminine to carry contraception (showing oneself to be a sexually active agent).

And those that are not:

(2a) Dorian chooses accountancy

(2b) Frieda chooses to use contraception, because she believes this is the best way to protect against pregnancy and sexually transmitted diseases.

Granting the substantive theorist the intuitions about the troubling cases, we should want to say that the first set of choice are troubling in a way that the second set are not. Thus, the account should be able to generate claims that respect this distinction. This is demanded, in particular, if we are concerned that those choices deemed to be autonomous are those entitled to the relevant normative benefits. I have noted, however,

that some theorists reject this demand on their account of autonomy. The claims in the following do not rely upon the acceptance of the relationship between autonomy and the normative roles it may play.

Recall that the substantive theorist might maintain that she is concerned not with this political and social aspect of autonomy, but rather with a notion more akin to personal authenticity, or some thicker sense of full self-determination. Even given this, we might nonetheless maintain that most choices are not inauthentic in a problematic way: Stoljar's intuition is that there is something distinctively problematic about the women's choices that are informed by oppressive norms. Most of our choices don't trigger such intuitions, and should not be diagnosed as problematic (non-autonomous, or inauthentic) by the substantive condition. But in any case to accommodate this concern, we can stipulate that the non-troubling choices, above, are authentic and fully self-determining. Dorian has been committed to accountancy all his life – he deeply believes in keeping the books straight and wants to ensure, in the small ways he can, that this is so. Frieda, likewise, is deeply committed to bodily integrity and control over reproductive choice. Each has been able to consider fully other options and values. These choices should be seen to be unproblematic, both in terms of autonomy, or full self-determination, or personal authenticity. We need to see, then, whether the substantive theorist is able, with the fully specified condition, to distinguish between choices of this kind and those that are intuitively troubling.

Can the substantive condition, once fully specified do this? Certainly, a substantive account may be able to explain why choices (1a) and (1b) are non-autonomous: indeed, it is with a focus on such cases that the conditions are formulated. Consider case (1a). Suppose that the value in question is that of self-respect. Lana might be committed to this value, and might believe her choice to be in accordance with it. But the substantive theorist will hold that the choice is not in fact in accordance with the relevant value, so the conditions set out in (HRC) are not all met.

Or, consider case (1b). Suppose the value in question is the true and good. Santa might value the true and the good, and believe her choice to be in accordance with this value, but insofar as she does not in fact choose in accordance with this value, she will not meet all the conditions in (HRC). Thus, each of these substantive accounts can, respectively, make sense of the claim that Santa's or Lana's choices are non-autonomous. (Note that, it may be that each of these conditions can make sense of both

cases: if the role of deference is deemed to be inconsistent with the good, or if the choice to take contraceptive risk is deemed to be inconsistent with self-respect, then one condition might be adequate to diagnose the problem with each – setting aside, for now the problems that I have claimed face each of these conditions.)

4.6.1 The problem of false negatives

However, insofar as the substantive condition is intended to make sense of apparently intuitive discriminations we make between autonomous and non-autonomous agents, (HRC) is inadequate for this task. This substantive condition is not, I will argue, able to classify the intuitively autonomous choices of Dorian and Frieda as autonomous. Thus (HRC) yields false negatives.

Suppose again that the value in question is the true and the good. Frieda might well be committed to this value, and might in fact choose in accordance with it. But even if we accept that the agent believes that her choice is in accordance with this value, it is not clear that this belief plays any role in her choice. Frieda's belief that using contraception is consistent with the good and non-oppressive might not play a role in her making this choice. Or suppose that the value in question is self-respect: Dorian may well be committed to this value, and she may in fact choose in accordance with it. But again, granting that he believes that the choice accords with the value, it is not clear that this belief plays any justificatory role in his choice: the belief that accountancy is consistent with self-respect might not play a role in his deliberations about what career to pursue. At best, then, the substantive theorist would have to reserve judgment about whether the choices of Dorian or Frieda are autonomous. But this is counter-intuitive; the kind of intuitive sway that the substantive theorists rely upon is precisely that, unlike run-of-the-mill – or even exemplary authentic, as we stipulated – choices such as Dorian's and Frieda's, there is something troubling about choices such as Santa and Lana's. The inability to make the kind of positive adjudications about such agents that are intuitive can be referred to as the problem of false negatives. With this problem the substantive account cannot say anything about what is held to be distinctively problematic about the choices at issue. Intuitions about the non-autonomy of the troubling choices are explained, but intuitions about the unproblematic choices are violated.

4.6.2 Response: counterfactuals

An obvious response is available to the substantive theorist. She might complain that we can make good sense of the justificatory role that the agent's belief about the consonance of her choice with the relevant value plays in bringing about her choice. All that is required in order to show that the agent's commitments and beliefs have an important role in bringing about the choice in such a way as to satisfy (HRC) is that the following counterfactual be met:

(cf) were the agent to believe that option O conflicts with value v, the agent would not choose it.

Taking this counterfactual into account (HRC) can classify agents such as Dorian and Frieda as autonomous in choice: their commitments and belief about the relevant value can be seen to play a tacit role insofar as were their respective choices not in accordance with the value, they would not so choose. This indicates a background policy, or filter, of some kind, according to which the commitments to the value and beliefs about the accordance of the choice with the value play an important filtering role. This condition takes into account the agent's tacit beliefs and the justificatory role they play in her actual choice. Plugged in, this gives us the following structure for substantive accounts:

- (HRC*): it is necessary for autonomous choice (jointly sufficient, with content-neutral conditions) that:
- a) the agent is committed to value v,
 - b) the agent believes that her choice is in accordance with this value, and
 - c) her choice in fact is in accordance with this value
 - d) her belief that the choice is in accordance with the value is (at least) part of the story for her so choosing, such that were she to believe otherwise – that is, believe that the choice conflicts with the value – she would not so choose.

Thus specified, the substantive account may be able to take into account the way in which the intuitively unproblematic choices of agents such as Dorian or Frieda, accord with the relevant value in a way that meets (HRC*), and so are autonomous. The problem of false negatives can be avoided, and the discriminations between the two kinds of choices can be made.

4.6.3 Counter-response: the problem of false negatives returns

Thus modified, however, many of the agent's choices of concern to the substantive theorist will meet this counterfactual epistemic condition. Consider the case in which

the agent is committed to value v , but that her choice fails to accord with it. For instance, Lana chooses to take contraceptive risk and is committed to the value of the true and the good. As Stoljar sets out the case, such an agent in fact chooses in accordance with the false and the oppressive – and because of the falsity of her beliefs, her autonomy is thwarted.

It is very plausible of such an agent that were she to believe that the choice conflicts with the value of the true and the good she would not so choose. Indeed, in claiming that the falsity of the norm is what prevents the agent from seeing how problematic it is, Stoljar at least implies that, were such an agent to have true beliefs, she would reject the guidance of the norm and choose differently.

Similarly for the agent who ‘misunderstands’ or fails to ‘effectively apprehend’ the value of self-respect. The problem, at least in part, as characterised by Hill, is that the agent does not have the right beliefs about the relation between the choice and the value. Were she to believe that the option conflicted with the value specified – as the counterfactual condition states – we can plausibly suppose that she would no longer choose it. With respect to this counterfactual, then, it is plausible that many of those agents that the substantive account has sought to classify as non-autonomous could in fact meet this condition.

Thus if the substantive theorist is to make sense of the intuitions that drive the positing of the condition, it is the world-guided condition that must do the diagnostic work. But in scrutinising (WRC), I claimed that we should reject the world-guided condition as necessary for autonomous choice. Insofar as the substantive theorist relies on the world-guided condition as a necessary condition, the problem of false negatives will remain. Recall Gus the Kantian and rude Betty; circumstances were such that Gus and Betty failed to choose in accordance with the specified values. Yet it is counter-intuitive to maintain that they are non-autonomous in their choice.

Misfortune and Rude are both cases in which changes in the world are such that, had the agent made the choice at a different time, or in a different context, under the same epistemic and deliberative conditions, the choice’s status would alter from autonomous to non-autonomous due to some change in the world. Cases in which mistakes are made due to changes in circumstance are surely pervasive – sufficiently pervasive that it is implausible to think that such a condition can be necessary for

autonomous choice. Once again, then, the problem of false negatives arises. Insofar as the condition is offered with the aim of making discriminations about which choices are and are not autonomous, such a condition seems unable to make the intuitive discriminations aimed for.²⁹

4.7 Summary and conclusions

I argued that work was required to get clear on exactly what the substantive condition demands. However, once the different readings of the condition are drawn out, we see that such a condition is unable to make intuitive discriminations between autonomous and non-autonomous choices. We should note that such a conclusion is not specific to the details of a particular substantive account that posits a particular value. For any value that is specified, it is implausible that an agent will meet the conditions set out by (HRC*) in very many cases. Thus substantive accounts face the problem of false negatives: such a view diagnoses many choices that are intuitively autonomous as non-autonomous. A substantive conception of autonomy is therefore ill-placed to capture the very intuitions that it is intended to make sense of.

Further, if a theorist does intend her account to figure in any of the normative frameworks, we should note how unacceptable this conclusion is: if an account of autonomy is to be well-placed to play the key normative roles of grounding respect, identifying the candidates for political participation, and protecting from paternalism, it should respect the intuition that, for the most part, our choices are autonomous and should not be intervened with. Insofar as a substantive condition counts a great many of our choices as non-autonomous, many deliberatively competent agents will fail to make choices that are protected from interventions, do not count in processes of collective

²⁹ It is worth noting that, were we to accept the world-guided condition as necessary (I have argued we should not) a further problem presents itself: it will sometimes – perhaps frequently – be very difficult to determine whether a choice does accord with a particular value. The complexity of the world means that sometimes – perhaps often – we are clueless about whether a certain value is consistent with a choice or not. Is the choice to eat organic consistent with the good? Or should one buy from fair-trade farmers? And so on. That this is so when the value in question is utility has been argued for at length by James Lenman. (See Lenman, J. (2000) ‘Consequentialism and Cluelessness’ *Philosophy and Public Affairs* 29(4), pp.342-370.) But we can well imagine that this will be so of other values: it will not be clear, due to what Lenman refers to as the ‘massive causal ramifications’ that all our actions have, whether a certain choice is in accordance with the value of the good, the non-oppressive, self-respect, and so on. Complexities in the world may not just obscure things for the agent deliberating *ex ante*, but for the theorist adjudicating *ex post*, about the autonomy of a choice. This makes the kind of intuitive discriminations the substantive theorist is aiming for even less likely, with the incorporation of a world-guided necessary condition.

decision making and so on. Thus whilst there may nonetheless be something problematic about those choices that are informed by oppressive values, we should not hold that this is a problem with the autonomy of the choice.³⁰

4.7.1 Further options for substantive conditions?

I have looked at, broadly, three versions of the condition: an epistemic version, world-guided version, and a reading that incorporates both epistemic and world-guided conditions. My argument against the incorporation of a substantive condition will only succeed as long as these readings are exhaustive. If there are other understandings of the substantive condition, then these might yet be adequate for making the distinction between intuitively non-autonomous and autonomous choices. What might further understandings look like? It is instructive, at this point, to note that the renderings of the condition I have set out mirror the literature that discusses whether the conditions for knowledge are internal or external.³¹ A brief look to the views in this debate will help us see what further options might remain for the substantive theorist.³²

Those who hold that the conditions for knowledge are internal hold that the justification for a belief must be accessible to the agent, upon reflection. Namely, she must have certain (second order) beliefs about her (first order beliefs about her) evidence for p, for her belief that p to amount to knowledge. Note the structural similarity of this view to that of (ERC), according to which an agent must have certain beliefs about her choice, for that choice to be autonomous. Externalists about knowledge, in contrast, hold that the agent need not have access to her justification for the belief that p – rather the justification for the belief can be external. If the agent consistently forms true beliefs (‘tracks the truth’), say, then insofar as p is in fact the case, the agent’s belief that p can count as knowledge. Note the structural similarity of this view to that of (WRC), according to which the agent’s choice is autonomous if it in

³⁰ In this, my conclusion is consonant with that of Clare Chambers, who remarks of such cases: “what is wrong is best perceived not as an issue of individual choice, flawed reasoning, or psychology but of the society in question” (Chambers, C. (2008) *Sex, Justice and Culture: The Limits of Choice*, Pennsylvania, Pennsylvania University Press, p.265). (Her concern is primarily with how liberalism might be equipped to criticise oppressive and unequal social arrangements, even when participation in those arrangements is (autonomously) chosen. Her conclusion is that liberals have the resources to criticise such inequalities and call for change in the social arrangements. However, she claims that the values that enable us to do so must be (primarily) equality, justice, and prevention of harm, rather than promotion of autonomy.)

³¹ I am grateful to Joe Morrison for encouraging me to consider these similarities.

³² In setting out these positions, I am drawing on characterisations from Pritchard, D. (2005) *Epistemic Luck*, Oxford, Oxford University Press.

fact accords with ('tracks') the value. Thus in (ERC) and (WRC) we have two starkly contrasting views, one internalist one externalist. In the epistemological debate we find a middle ground, in which we find views that focus on the agent's reliability or sensitivity.

Might an analogue to these middle ground views be available to the substantive theorist? Let us consider a sensitivity view, which holds that the agent's belief that *p* counts as knowledge if it is sufficiently sensitive to changes in the world (understood in terms of whether, in near possible worlds in which *p* is false, the agent does not believe that *p*). Could substantive constraints be built into the conditions for autonomous choice by incorporating a sensitivity condition?

Views that bear structural similarity to these sensitivity accounts are to be found in the literature on autonomy and self-governance. Indeed, Fischer and Ravizza hold that an agent can be understood to be 'reasons-responsive' in her choice, *c*, if, in a possible world in which there is sufficient reason to choose otherwise, the agent does not choose *c*.³³ Might such a view be of use to the substantive theorist? She could modify her condition to hold that:

(R-RRC) an agent's choice is autonomous if an agent has proper appreciation of the value *v*, where 'proper appreciation' requires that, in some possible world in which there is sufficient reason to choose in accordance with value *v*, the agent so chooses.

Such a condition captures the thought that the agent can be sensitive to what is important about the value *v*, although she may not always choose in accordance with it – due to overriding factors, or having made a mistake, and so on.

Likewise, a 'reliability' view focuses on the agent's general reliability, over time, in forming true beliefs (though on occasion, she may make mistakes). If an agent is sufficiently reliable, her true beliefs constitute knowledge (though not on the occasion that she has made a mistake). An analogous view may focus on the agent's reliability, over time, in acting in accordance with the specified value (even though on occasion, she might fail to do so). If she is suitably reliable, her choices can be regarded as autonomous (were the analogy strict, we might have to claim that, on occasions on

³³ See Fischer, J. & Ravizza, M. (1998) *Responsibility and Control: A Theory of Moral Responsibility*, New York, Cambridge University Press, esp. at Chapter 2 and 3.

which she has made a mistake, she is non-autonomous in choice. But given that autonomous choice is not factive in the way that knowledge is, this pressure is somewhat dispelled).

If the substantive theorist adopts such a sensitivity or reliability condition, however, they have given up a great deal. Indeed, the main requirement of the substantive condition, (RelChoice) is dropped: a condition for autonomy is no longer that the agent must choose in accordance with the value *v*. Rather, the ‘normative substance’ of the account is built in by placing demands on the agent’s capacities – their sensitivity to the value, or reliability in choosing and acting in accordance with it. The focus, then, is shifted from the agent’s choice onto the agent’s capacities; what matters is not whether or not the agent’s choice is in accordance with the value, but rather whether or not the agent could choose in accordance with it, or has the capacity to appreciate the value. Such a view would be similar to that described by Benson, which “incorporates normative content but does not constrain directly the types of actions agents might autonomously perform, or the content of the motives or values that lead them to act”.³⁴

Once we move away from the demand that the choice accords with the value, towards a more nuanced incorporation of demands for appreciation of the value, we move towards a distinctive kind of view. Such a view does not substantively constrain the objects of autonomous choice. Rather, it focuses on the competence of the agent: such views are concerned with the conditions for autonomous agency. In the next part of the thesis, I consider some such views, and their attempts to incorporate relational conditions for autonomous agency.

The conclusion of this part of the thesis is thus a negative one. We can leave behind the attempt to incorporate relational content by positing substantive conditions for autonomous choice – I have argued that such conditions are unmotivated and cannot serve the purposes for which they are intended. Further, if we seek an account that can play a role in the normative frameworks that account for autonomy’s value, we should not accept one that incorporates value-laden conditions. In the next part of the thesis, then, I look at an alternative way to incorporate relational conditions into a conception of autonomy: by offering relational conditions for autonomous agency.

³⁴ Benson, P. (2005b) ‘Feminist Intuitions and the Normative Substance of Autonomy’ in Taylor, J.S. (ed.) *Personal Autonomy*, Cambridge University Press, pp.124-142. See p.125.

Chapter 5. Relational accounts of autonomous agency: Benson and Oshana

This part of the thesis is concerned with the conditions autonomous agency. Rather than specific choices, we are here concerned with the capacities that an agent has and her ability to exercise them on a range of occasions. A commonsense understanding of autonomous agency can be stated as:

(AGENCY) an agent is autonomous when she has the capacities, and the ability to exercise the capacities, that are necessary for self-governance of her deliberation, choice and action.

In this chapter I will consider two ways of developing this commonsense understanding. First, I will look at whether we can understand the account from Paul Benson as one that offers constitutively relational conditions (section 5.1 – 5.4). Whilst I will argue that his account should not be understood as offering constitutively relational conditions, there are nonetheless important points to draw from examining his view. I will then turn to Marina Oshana's relational account (section 5.5). Her account shows one clear and distinctive way of building constitutively relational conditions into an account of autonomous agency. However, I will argue in this chapter that her constitutively relational condition is unmotivated on two counts.

5.1 Benson's account of autonomous agency

Like Stoljar, Benson motivates his condition for autonomous agency by pointing to agents who intuitively lack autonomy, but with whom the content-neutral conceptions (of the kind I outlined in Chapter 2, above) find no fault. Looking to the interaction between the agent and the social relations in which she stands, Benson claims, can help us make good sense of the non-autonomy of agents in such instances. It looks like this kind of view might be a good candidate for a constitutively relational account of autonomy. I'll examine whether this is so.

First, I set out a case that motivates Benson's relational condition. Then I will consider the way in which Benson's account might be understood as incorporating a constitutively relational condition. However, we will see that the most plausible

rendering of the condition is as causally necessary. It is instructive to consider this account, however, as two important points come to light: first, that an account can incorporate some values, or ‘normative substance’ into the account without substantively constraining the objects of autonomous choice (as those accounts considered in the previous part of the thesis do). Second, we see that an account that posits causally, rather than constitutively, relational conditions for autonomy can nonetheless be well positioned to take into account autonomy’s social aspects. An account that remains focused on the psychological states of the agent need not appeal to

the imagery of the inner citadel ... [which] suggests a picture of autonomy as relegated to the background of social life – a characteristic that emerges behind an invisible partition that isolates each individual from the rest, overlooking entirely the social and relational dimensions of self-government.¹

This charge comes from Oshana, but we shall see it to be unfounded.

5.1.1 Benson on ‘gaslighting’

Benson argues that content neutral accounts are unable to make sense of certain cases that appear to be clear failures of autonomous agency. A central case in Benson’s argument is:

Gaslighted² [America, the end of the 19th century]: Ingrid is diagnosed with hysteria by her well meaning but mistaken husband, who is also a doctor. Ingrid does not dispute or resist her husband’s diagnosis; she believes his judgement is informed by and made “on the basis of reasons that are accepted by a scientific establishment which is socially validated and which she trusts”. As a result, Ingrid “ends up isolated, and feeling rather crazy”. Although her competences in deliberation are perfectly adequate, Ingrid no longer trusts her competences as an agent; in particular, her competence to respond to reasons – “she has ceased to trust herself to govern her conduct competently”.³

¹ Oshana, M. (2006) *Op.Cit.* p.51.

² Gaslight: “to manipulate (a person) by psychological means into questioning his or her own sanity” (<http://www.oed.com/>). Benson’s example is a significant reworking of the plot in the 1944 film *Gaslight* (in the film the gaslighted is not well-intentioned).

³ Benson, P. (1994) *Op.Cit.* See pages 655-657 for his description of this case. The unquoted text is a paraphrasing of his description.

This case informs the condition for autonomous agency that Benson sets out in his (1994) paper, and builds on in later work (in particular, in his 2000, 2005a and 2005b⁴). An agent such as Ingrid, Benson claims, fails to be an autonomous agent. Nonetheless, she is able to meet a range of content-neutral conditions for autonomous agency. It is plausible that Ingrid will meet the conditions set out by structural content-neutral accounts of autonomy. For example, Ingrid retains reflective competence, and a sense of her values, and may retain executive and regulative power over her will; moreover, it is plausible to suppose that Ingrid will endorse the way in which she came to form this attitude towards herself (on the basis of advice from her trusted doctor and husband).⁵

The problem with a character such as Ingrid cannot be explained by the existing content-neutral accounts, Benson claims. Rather, despite the adequate functioning of her autonomy-relevant capacities for reflection and deliberation, the agent lacks a sense of her own competence in the exercise of them, and her worthiness to do so. As a result, even when she puts these capacities to use, “she is quite disengaged from her actions”.⁶ When she chooses and acts, she doesn’t trust that she is doing so competently, and feels unsure that she could account for her actions to others. This disengagement is a defect that undermines autonomous agency, Benson claims. We can now turn in more detail to Benson’s characterisation of such a failure, and the condition for autonomy that he sets out accordingly.

5.1.2 Benson’s condition

Although Ingrid’s capacities for autonomous agency are intact, Benson claims that as a result of her lack of confidence in her abilities, Ingrid’s “identification with the possession and exercise of those powers [of reflection on and regulation of action]... has been threatened by her revised view of her own competence.”⁷ She does not trust her abilities, even though they are in fact fully functional.

⁴ See Benson, P. (2005a) ‘Taking Ownership: Authority and Voice in Autonomous Agency’ in Christman, J. and Anderson, J. (eds.) 2005, Oxford, Oxford University Press, pp.101-126; see also Benson, P. (2005b) *Op.Cit.*, and Benson, P. (2000) *Op.Cit.*

⁵ See chapter 2 for overview of structural content-neutral conditions. In particular, I am here referring to a coherence condition, a reflective competence condition, and a historical condition. See Christman, J. (1991b) *Op.Cit.*; See Watson, G. (1975) *Op.Cit.*; See Frankfurt, H. (1971), *Op.Cit.*; respectively, and Christman, J. (1991b) again for a historical condition. Note that, in his (2005a), Benson rejects some of the content-neutral conditions considered in chapter 2; namely, those Frankfurtian endorsement conditions that make ambivalence autonomy undermining.

⁶ Benson, P. (1994) *Op.Cit* p.656.

⁷ *Ibid*, p.657.

The thought is that an autonomous agent must have a sense of her own competence as an agent, and this requires having a sense of one's competence to grasp and engage with reasons. The attitude in particular that Benson identifies as incompatible with autonomous agency, insofar as it impairs identification with one's capacities, is the agent's sense of being "insufficiently worthy to relate with others because of the normative expectations that the agent judges to apply to those relations".⁸ Gaslighted Ingrid, taking herself to be unhinged, does not feel she can engage in relations with others according to the norms she believes to govern those relations, Benson claims. Put briefly, she does not believe that she can engage in social relations in which others hold her as competent to account for her actions. This is because, having been made to 'feel crazy', she does not trust her grasp of the standards that others hold her to, or her competence to meet those standards, or to engage with others. Accordingly, Benson claims that:

the sense of worthiness to act which is necessary for free agency involves regarding oneself as being competent to answer for one's conduct in light of normative demands that, from one's point of view others might appropriately apply to one's actions.⁹

The point here is put in terms of the agent's sense of her worthiness. Later, Benson talks of the agent's sense of her own authority. Autonomous agents possess, he claims,

attitudes towards their own socially situated authority to construct, stand by and speak for their reasons for acting.¹⁰

The two relevant attitudes – of authority and worthiness – are connected in the following way:

standards for agents' authority to construct and potentially answer for their reasons for acting enter into autonomy by way of the attitudes toward their own competence and worth through which agent's claim such authority.¹¹

On Benson's later view, then, the thought is that a necessary condition for autonomous agency is that the agent takes herself to have authority to account for her actions. Having this kind of authority requires that the agent has a sense of her own competence

⁸ *Ibid*, p.660.

⁹ *Ibid*, p.660.

¹⁰ Benson, P. (2005b) *Op.Cit.* p.133.

¹¹ *Ibid*, p.136.

and worth; this competence and worth grounds her claim to such authority. Indeed, on his later view (2005a) Benson emphasises that an agent must actively take authority:

we should conceive of this authority as depending, in part, upon an active process of authorisation that autonomous agents enact upon themselves.¹²

Authorising oneself in this way, Benson makes clear, can amount to having a certain attitude towards oneself. It might also amount to explicit vocalisations – we can imagine Ingrid reassuring herself “I do have my own reasons. I can address others on equal terms, and give them these reasons” – but it need not. Rather it requires having a certain kind of self-regard: a sense of one’s competence, and worth or authority to speak for one’s actions. Thus Benson demands, in addition to the requisite content-neutral conditions, a condition, of the form:

(RelAgency1) A necessary condition for autonomous agency is that the agent has authority as an accountable agent; that is, she holds herself to be competent to answer for her actions (according to the normative standards of the social context).

Part of the project of this thesis, recall, is to consider what a constitutively relational condition for autonomy might look like. The task of the following section will be to consider whether we should understand a condition such as (RelAgency1) as a constitutively relational condition for autonomous agency. I will argue that it is most plausibly understood as attending to the potential causal impact of social relations on autonomous agency.

5.1.3 Benson and constitutively relational conditions

When first setting out his account, in his (1994) paper, it certainly looks as if Benson is intending to offer a constitutively relational condition for autonomous agency. He writes that his view goes beyond those in which “the character of our interpersonal or social situation, or of our conception of it, can only affect our free [autonomous] agency accidentally, by virtue of its potential influence on our capacities to express what most matters to us in what we do”.¹³

¹² Benson, P. (2005a) *Op. Cit*, p.114.

¹³ Benson, P. (1994) *Op. Cit*, p.666.

If relational conditions are causal, then the impact of social relations on autonomous agency can be merely ‘accidental’; certain social relations may or may not cause an agent to lose her sense of competence, her sense of her authority to account for her actions. But if relational conditions are constitutive of autonomy, then their impact will not be merely contingent. The failure of certain social relations to obtain just will be a failure of autonomous agency; for the agent who fails to stand in those social relations fails to be autonomous. Given this, then, there is reason to suppose that Benson is aiming for a constitutively relational understanding of autonomy. Indeed, in his later version, which focuses on the agent’s attitude of self-authorisation, Benson writes, of content-neutral conditions that:

they do not recognise any inherent, constitutive connection between agential ownership and persons’ social relations. They entail the notion that persons can own their motives independently of their socially structured authority to stand by what they do. My proposal contrasts with the constitutive individualism of other theories.¹⁴

Again, whilst not explicit, this is suggestive of the claim that Benson is offering a constitutively relational account. But what would this, in fact, demand for autonomous agency? What social relations, on Benson’s view, would be constitutive of autonomy?

Benson focuses on the agent’s authority to account for her actions. Is standing in social relations in which an agent is regarded as authoritative in this way necessary for autonomy? No: Benson holds rather that it is the agent’s attitude towards her competence and worthiness that grounds her authority. Indeed, he is emphatic that “agents’ authority arises through their self-authorisation”.¹⁵ It is not necessary that the agent is regarded as authoritative by others – although such a failure of regard may eventually erode the agent’s sense of her own authority, as he acknowledges.

Might we say, then, that agents ought to stand in social relations in which they are regarded as competent and worthy in order to secure autonomous agency? Again, Benson himself claims that we should not see such relations as constitutive of autonomous agency:

that interpretation seems to rule out the freedom of any persons who are not aware of, or committed to, the conventional social norms by which others are likely to assess their conduct [hence to whom the agent does not feel

¹⁴ Benson, P. (2005a) *Op.Cit.* p.108.

¹⁵ *Ibid*, p.110.

competent to answer to]. Likewise, it might seem to preclude the freedom of any who have tried to separate themselves from most common forms of human community and social practice.¹⁶

Insofar as an agent repudiates convention or society, she may not stand in social relations in which she is regarded as competent or worthy to answer to others (either because she is not regarded as competent, or because she does not stand in any social relations, so is not held by others in any regard). But, Benson writes:

Such agents lack the knowledge, commitments, or bonds that are prerequisites of their being answerable to others for their actions ... but this in no way has to threaten their freedom [autonomy].¹⁷

Rather than claim that others must regard the agent as competent, Benson emphasises that it is the agent's 'subjective sense of normative competence' that is required for autonomous agency. It is consistent with having such a sense of competence and worth that an agent cannot, in fact, answer for her actions; nor need she be regarded by others as competent.

It is clear, then, that Benson's view should not be understood as a constitutively relational account of autonomous agency. Whilst he attends to the role that interpersonal relationships play in impacting upon the agent's sense of competence and her authority to answer for her actions, we should not infer from this that there are constitutively relational necessary conditions for autonomous agency. It is not necessary that an agent stand in relations in which she is regarded as authoritative; nor is it necessary that an agent stand in relations in which she is regarded as competent. Such social relations may, contingently, impact upon an agent's sense of competence; but an agent may have, and be quite confident in, her capacities for autonomous agency, and regard herself as authoritative to account for her actions, whilst failing to stand in such social relations (or indeed any social relations at all). We can accept (RelAgency1), and acknowledge that Ingrid's autonomy is undermined, in gaslighted; but this is not due to her failing to stand in some social relations that are constitutive of autonomy. Rather it is due to the causal impact of the social relations she occupies upon the self-regarding attitudes necessary for autonomy.

¹⁶ Benson, P. (1994) *Op.Cit.* at p. 661.

¹⁷ *Ibid*, p.662.

5.1.4 Normative content and social sensitivity

There are two important points to draw from Benson's account. The first point harkens back to the issues considered in the previous three chapters. There I considered attempts to offer value-laden conditions for autonomy. With Benson's account, we see that one can incorporate 'normative substance', as he puts it, into an account of autonomy without substantively constraining the contents of autonomous choice. Rather, Benson builds in some normative substance with his demand that the agent has a sense of worthiness and competence in her ability to account for her actions. Thus he acknowledges the value of an agent's sense of competence and worthiness to account for her action.

However, his account does not simply demand that an agent must choose in accordance with the value of her own agency, or some other value (such as those considered in chapters 2-4). Rather, Benson brings to light that the value of this sense of competence is an enabling one: without such self-authorisation, agents are alienated from their agency. But incorporating value or 'normative content' in this way does not require restricting autonomy such that it is consistent only with those choices that accord with particular values. An agent may fail to choose in accordance with some specified value, but nonetheless remain sure of her competence and worthiness to account for her action to others.

As we have seen, Benson rejects the intuition, from Stoljar, that agents who choose in accordance with false and oppressive norms are non-autonomous (and I think he is right to do so). With his condition, we are able to acknowledge that agents who act for reasons we consider to be bad reasons – such as those based on false or oppressive norms – can nonetheless be autonomous agents insofar as they meet any required procedural conditions and have the relevant sense of worth, meeting (RelAgency1).

The second point concerns the worry raised at the beginning of this chapter. Benson's account focuses on the psychological states of the agent; namely, the (content-neutral) procedures she undertakes in deliberation, as well as her attitude towards her competence to engage in such procedures. This view, then, falls into the class of views that focus on the psychological states of the agent. Oshana in part motivates her constitutively relational view (to be considered shortly) with claims about the inadequacy of views which focus on the psychological states of the agent. She claims

that these views ‘cordon off’ the notion of autonomy from the social relations in which the agent stands:

the emphasis of [the psychological state] accounts on the subjective point of view [of the agent] reflects the belief that safeguarding the autonomy of persons consists in safeguarding what is symbolically described as the inner citadel.¹⁸

We saw her concern that focusing on such an ‘inner citadel’:

suggests a picture of autonomy as relegated to the background of social life – a characteristic that emerges behind an invisible partition that isolates each individual from the rest, overlooking entirely the social and relational dimensions of self-government.¹⁹

She goes on to express the concern that such a view “ignores the socially embedded character of human identity”.²⁰ Having considered Benson’s view, which focuses on the psychological states of the autonomous agent as conditions for autonomy, we can now see that such concerns are unfounded. It is worth briefly considering each of Oshana’s concerns in turn.

First, the focus on the agent’s sense of competence and worth to account for her actions means that securing autonomy involves (amongst other things) that she secures a certain attitude. But this does not preclude attention to social relations: indeed, insofar as certain social relations may erode the required sense of competence (as Benson’s example indicates), his view enables the identification of the kinds of social relations it is worth securing if one cares about fostering autonomy: namely, social relations in which agents hold each other in mutual regard as competent and accountable agents.

Second, insofar as the attitude that Benson identifies is one which, as he notes is “sensitive to others’ attitudes towards the agent”²¹ there is no reason to suppose that an account of autonomy that attends to the agent’s psychological states is one which isolates individuals and overlooks the relational elements of self-governance.

¹⁸ Oshana, M. (2006) *Op.Cit.* p.51.

¹⁹ *Ibid*, p.51.

²⁰ *Ibid*, p.51.

²¹ Benson, P. (1994) *Op.Cit.* p. 659.

Finally, Benson's focus on the sensitivity of agent's sense of competence to others attitudes acknowledges that core mental states of the agent are affected by the situation in which they are embedded. Indeed, he emphasises that persons are social beings, noting that by highlighting the relevance of social relations to self-governance, we can explain why individuals "have reason to care about maintaining their status as full [i.e. competent and accountable] participants in social relationships and practices".²² Insofar as failing to occupy such relations runs the risk of undermining the attitudes integral to autonomy, agents have an important stake in relations in which they stand. Benson concludes that "both the capabilities and attitudes this position demands concern interpersonal exchange... There is nothing unduly individualistic about the conception of selfhood this view might suggest".²³

Focusing solely on the psychological states of the agent, then, need not entail ignoring the social and relational aspects of autonomous agency. Such aspects may have an important causal impact upon autonomous agency. Thus Oshana's concerns can be addressed and rejected. Such concerns motivate Oshana's account, which demand that the agent occupy certain social relations in order to be autonomous. Her account, as we shall see, clearly gives relational conditions a constitutive role. I turn to consider it now; whilst one motivation for such an account has fallen away, it might nonetheless be true that an agent must occupy the relations she posits. However, I will argue in the next section that there is no good reason to suppose this is the case. Then, in the next chapter, I will argue that there is good reason to hold that it is not the case.

5.2 Oshana's account of autonomous agency

Insofar as social relations are incorporated into an account in terms of the impact they may have on the agent's psychological state, the account is not constitutively relational. This is because it is plausible that the relevant psychological state can be held in the absence of social relations that foster it: the agent may be particularly robust; or the psychological state could be induced in some other way.

Oshana's account, however, has a structure that is clearly constitutively relational. She demands that, in order to be an autonomous agent, an agent must stand in certain social relations. The central thought is that not only must the agent's

²² *Ibid*, p. 667.

²³ Benson, P. (2005a) *Op. Cit*, pp.118-119.

psychological states be a certain way; the world must also be a certain way.²⁴ Thus her condition is ‘externalist’, demanding as necessary for autonomy certain states external to the agent. The condition that Oshana argues for is:

(RelAgency2) A necessary condition for autonomous agency is that the agent stand in social relations in which she has de facto power to choose and manage her life; in which her choices are from an adequate range of options; in which she has power counterfactually, such that she is able to maintain control over her life even in possible scenarios in which others attempt to intervene.²⁵

In the rest of the chapter, I will examine Oshana’s argument for this condition, and note that there are some crucial premises that remain unmotivated. The preliminary conclusion of this chapter is that the considerations that allegedly push towards the adoption of a constitutively relational condition need not be heeded; there are no good reasons for adopting a constitutively relational condition. In the next chapter, I will argue that important considerations counsel against demanding that an agent occupy certain social relations in order to be autonomous.

5.2.1. General intuitions

Oshana argues for her relational condition by setting out some broad intuitions about autonomy, and then some cases in which, according to these intuitions, the agent is non-autonomous. She then argues that her relational condition can best explain the intuitions about the particular cases, as well as the broader intuitions that she sets out. Here, then, are the general intuitions about autonomy that Oshana takes as a starting point:²⁶

²⁴ I focus here on the social relations that are claimed to be necessary for autonomy. This is likewise the main focus of Oshana’s discussion. But she also maintains that other natural phenomena may prevent autonomy from obtaining, insofar as they prevent the condition she sets out from being met: “certain natural phenomena – physical ability, for example – do not simply condition the exercise of global autonomy; by increasing or decreasing ‘the range over which, or the ease with which, it can be enjoyed,’ these phenomena yield autonomy of a more resilient or more enfeebled form” (2006, *Op.Cit.*, at p.92).

²⁵ These conditions are set out in Oshana, M. (2006) *Op Cit.* pp. 83-90. Her setting out of social-relational conditions here builds on her (1998) account of relational autonomy. See Oshana, M. (1998), ‘Personal Autonomy and Society’ *The Journal of Social Philosophy* 29 (1), pp. 81-102.

²⁶ In her (1998), Oshana sets out a different set of motivating intuitions:

- a. The autonomous agent is self-directed
- b. the autonomous agent is in control of her choices, actions, will
- c. the autonomous agent is in control of her external circumstances
- d. the autonomous agent can pursue, effectively, her choices or goals in action
- e. The autonomous agent is not subject to psychological or physical interferences (compulsions, constraint, coercion etc)
- f. Autonomy is “the good which paternalism fails to respect”

A) The control intuition: “Autonomous beings are in actual control of their own choice, actions and goals.”²⁷

This is plausible. Oshana goes on to qualify the control intuition: “By this I mean that the person is in possession of the de facto power to govern herself”.²⁸ By de facto power, Oshana means that the agent must have the power to decide what to do, from a range of live options. In addition, having this power involves that:

- the agent face minimal interference
- the autonomous agent acts for her own reasons
- the autonomous agent has an adequate degree of self-control.²⁹

These latter three elements are relatively uncontroversial, and I will not comment upon them further here.

B) The authority intuition: “Autonomous persons are in a kind of authoritative control of their own choices, actions and goals”.³⁰

This kind of control, Oshana claims, requires that an agent ‘own’ the management of her choices (in the sense that Benson demands, as discussed above). Moreover, it requires that the agent “not only have de jure control over her choices, actions, and goals but must enjoy a status against other persons or institutions ... having the relevant kind of authority guarantees that a person’s life is free from the domination of others.”³¹ The thought is that this kind of domination undermines autonomy by preventing the agent from having de facto power: insofar as another agent could intervene arbitrarily, even when she does not, she has “unwarranted power over him. He cannot govern himself”.³²

I paraphrase from Oshana, M. (1998) *Op.Cit.* esp. pp.81-82. Note that these intuitions, without modification, are too strong: intuition c cannot be satisfied by the most self-directing agent. Intuition f is one that I focus upon in more detail in the next chapter. The intuitions stated here cover a lot of ground, and one might wonder whether Oshana is concerned with a range of aspects of autonomy: choice, agency and action. The intuitions she sets out in her (2006) mark a significant refinement of these, so I focus on her later claims only.

²⁷ Oshana, M. (2006) *Op.Cit.* p. 3.

²⁸ *Ibid*, p.3.

²⁹ I paraphrase from Oshana, M. (2006) at pp. 3-4.

³⁰ *Ibid*, p. 4.

³¹ *Ibid*, p.4.

³² *Ibid*, p.90.

The two intuitions from which Oshana starts, then, hold that the agent must have control over her choices and actions, and that the agent must be authoritative over them. In cashing these out, we see that Oshana has in mind that a certain kind of control and a certain kind of authority are required: de facto control, and non-dominated authority. Oshana intends these intuitions to be ones that “are not controversial” and claims that “they are ideas held ... by laypersons and academic philosophers alike”.³³ I will show that they are in fact controversial, and that weaker understandings of both of these conditions are available to us: these weaker understandings do not require constitutively relational conditions for autonomy.

Before looking in more detail at the examples, it is worth marshalling the following concern: even at a first glance, the condition that Oshana sets out and the intuitions that drive it look to be concerned with a different sense of autonomy than that at issue in the previous chapters. ‘Autonomy’ is sometimes used to refer to a particularly valuable way of life – a life with a good range of options, say. But this is not the sense with which we have thus far been concerned; rather, we have been concerned with the nature of the choices, and the kind of capacities, that agents must have if they are entitled to the normative benefits of respect, protection from paternalism, and so on. It might be suggested that Oshana is concerned with the former sense of autonomy, rather than the latter (an objection she refers to as “the apples and oranges objection”: both are fruit, but they are different kinds of fruit).

However, whilst at some points Oshana appears to talk of a different sense of autonomy, she explicitly states that her concern is with the same sense of autonomy that we are here concerned with, and that her account is offered as a rival account to those that focus on the psychological states of the agent, such as Benson’s:

this [social relational account] is not autonomy of a different sort than the autonomy that the psychological authenticity accounts analyse. The advocate of autonomy as psychological authenticity is not attempting to explain something other than the autonomy of persons ... in all instances the object of concern is the status of persons.³⁴

It is clear, then, that Oshana intends to be offering conditions that fill out the same sense of autonomy as that with which Benson, and this thesis, is concerned.

³³ *Ibid*, pp.3-4.

³⁴ *Ibid*, pp. 93-94 (my emphasis).

5.2.2 The examples

Oshana argues for her relational condition by way of example. She claims that, in accordance with the general intuitions she sets out, the agents in the examples are intuitively non-autonomous. Then she argues that only her constitutively relational conditions for autonomous agency can make sense of the non-autonomy of the agents. I will first set the key examples out, showing how they are intended to support Oshana's relational account. Then I will tease out the different understandings of the control intuition that we might adopt. I will show that weaker versions of the intuitions stated above are available to us, and deliver starkly contrasting verdicts about the example cases, and about the need for constitutively relational conditions for autonomous agency.

The deferential wife

Oshana imagines a woman, named 'Harriet', who has chosen the role of deferential homemaker and wife. Harriet meets all the procedural conditions that might be required for autonomous agency:

let us assume as well that Harriet finds her life gratifying and has no wish to alter it. There is nothing she values more or wants more than to be the angel in the house ... she possesses all the autonomy competences.³⁵

Despite meeting the demands for reflection and deliberation upon her preferences and their origin, and having the relevant capacities for autonomous agency, Oshana holds that nonetheless such an agent fails to be autonomous. Whilst this claim is in line with the view of Thomas Hill Jnr., considered in Chapter 3, Oshana rejects the thought that the deferential individual's non-autonomy is to be explained by reference to the contents of her choice:

the substantive element of social-relational autonomy is not generally found in the object of the person's desires, and in any event Harriet's lack of autonomy is not due to her lamentable desires ... her lack of autonomy is due to her personal relations with others and to the social institutions of her society.³⁶

³⁵ *Ibid*, pp. 58-59. The phrase 'angel in the house' is from Woolf, V. 'Professions for Women' a speech before the National Society for Women's Service, (1931), in Woolf, V. *The Virginia Woolf Reader*, Leaska, M. (ed.). (1984), at p.278.

³⁶ Oshana, M. (2006) *Op.Cit.* p.59

In particular, the social relations that thwart her autonomy are ones in which “afford her less recognition and independence than she deserves and than she might otherwise have”.³⁷ She goes on:

Let us assume that the social relations Harriet is party to, given her role as homemaker, afford her less financial flexibility, less social mobility, and fewer opportunities for intellectual and creative development than she could have were these relations otherwise. Suppose, too, that in her relationship with her spouse she makes none of the important decisions ... Absent too are economic and political institutions that might empower homemakers.³⁸

Why is it that social relations such as these prevent an agent from being autonomous? These relations and institutions prevent the agent from having the right kind of control over her choices and actions. Focusing in on a particular choice may be helpful here. Let us suppose that Harriet chooses, in deference to her husband and in accordance with his wishes, to take a holiday in Luxembourg. She prefers this option also, and is perfectly happy with the choice. But in so choosing, Harriet lacks *de facto* control of her choice: holidaying in Luxembourg was in fact the only option she had available to her. This, along with the fact that she is not afforded recognition and independence, means that she lacks counterfactual power – had she wished to holiday in Lebanon instead, she could not have so chosen. Were her husband to attempt exert control over her and interfere with her choice (forcing her to abide by his wishes for a trip to Luxembourg rather than Lebanon) she would not be able to maintain control over her choices.

Thus the social relations in which she stands are such as to ensure that the agent fails to meet (RelAgency2). That condition, recall, demanded that:

(RelAgency2) A necessary condition for autonomous agency is that the agent stand in social relations in which she has *de facto* power to choose and manage her life; in which her choices are from an adequate range of options; in which she has power counterfactually, such that she is able to maintain control over her life even in possible scenarios in which others attempt to intervene.

Harriet lacks *de facto* power, and lacks the power to maintain control over her life in the instance that her husband (or perhaps another) were to attempt to intervene. Recall that Harriet meets all the procedural and content-neutral conditions. Thus to explain the

³⁷ *Ibid.*

³⁸ *Ibid.*, pp.58-59.

intuition that she is non-autonomous, Oshana claims, we must appeal to the relational condition (RelAgency2): it is only because Harriet fails to stand in social relations – that ensure her recognition and independence that would secure power and authority – that she lacks autonomy.

The monk

Suppose a monk chooses and occupies his role with all the competences required for autonomy. Whilst in his role as a religious devotee, “the religious order has power over him sufficient to compel him to behave in a certain way”³⁹ (although they may never exercise this power). However, the monk is able to decide, on a yearly basis, whether or not to remain in this role. Such an individual, Oshana claims, is non-autonomous. Once again, this cannot be explained by conditions that pertain to his psychological states, or by reference to the substantive content of his choice. Rather, the social conditions that he occupies are such as to prevent him from being autonomous. In his role of religious devotee, he lacks de facto power: “his monastic superiors preserve authority in the interim, for his life is ruled by them on a daily basis”.⁴⁰ Again, it will be useful to focus on a particular choice. Suppose at the beginning of the week, the monk chooses, in accordance with monastic practice, to undertake a vow of silence. The monk lacks de facto power, because this is the only option that he has. Were he to choose otherwise (to chatter away to himself and to attempt to with others) his choice would not be effective – we can suppose that he would be pulled up and disciplined by the monastic superiors – this other option is not a ‘live’ one. Moreover, were he to attempt to choose otherwise, he would lack the authority to prevent others from intervening: the monk is not in a position to maintain control over his choices should the monastic superiors attempt to intervene.

Thus the social relations in which the monk stands, whilst in the monastic order, prevent him from achieving autonomy, on Oshana’s view. She notes that we should concede that the monk has ‘potential autonomy’, due to the fact that, annually, he has the authority to leave the monastery, and reinstate himself in social relations in which he has de facto power and authority. But “the fact that the monk can annul his status as non-autonomous does not mean he is self-governing”⁴¹ Oshana claims. Note that competences of the agent remain the same whether he is in the monk role or not. So the

³⁹ *Ibid*, p.62.

⁴⁰ *Ibid*, p.63.

⁴¹ *Ibid*, p.63.

intuition that the monk is non-autonomous whilst in the role of religious devotee can only be explained by the relational condition: lack of actual power and authority is what explains the non-autonomy of such an individual, Oshana claims.

Oshana's argument, then, can be understood as follows: in accordance with the broader intuitions (about control and authority) the deferential wife and the monk are intuitively non-autonomous. The content-neutral conditions that focus on the psychological states of the agent cannot make sense of their non-autonomy. Both individuals lack de facto power and the authority to maintain control if others were to attempt to intervene. Having such power and authority requires that the agent stand in certain social relations. Thus relational conditions – those which shore up the power and authority that these individuals lack – are necessary for autonomous agency.

In the following section, I reconstruct the argument in more detail. In doing so, we will see more clearly the role that the control and authority intuitions are playing. I will then tease out a different understanding of the control intuition. It is not uncontroversial to accept the stronger understanding of the control intuition that informs Oshana's judgments about the non-autonomy of the deferential wife and the monk. I will argue that the stronger understanding that she endorses lacks independent motivation.

5.2.3 Examining the argument

There are two parts to Oshana's argument: first, the claim that the agents in her examples are non-autonomous, as informed by the 'uncontroversial' intuitions; second, the claim that the non-autonomy of the agents cannot be explained without her constitutively relational condition. We have already seen, in earlier chapters, that it is not uncontroversial to hold that agents such as the deferential wife are non-autonomous, and that it is difficult to substantiate such a claim. I will here argue that the more general intuition that informs Oshana's judgment on this is also controversial.

A reconstruction:

More formally, Oshana's argument appears to have this form:

- 1) autonomous agents have control over their choices, will, and action (the control intuition)

- 2) certain roles (e.g. a deferential role) are incompatible with the agent having this kind of control over her choices, will and action.
- 3) such roles (e.g. deferential roles) are incompatible with autonomous agency.

This is the first part of the argument. From there, Oshana goes on to explain that it is because the relational conditions fail to obtain that the agent is non-autonomous:

- 4) these roles are compatible with the internalist, psychological conditions being met.
- 5) any failure of autonomy must therefore be because the required relational, or 'externalist' conditions fail to obtain (that is, the conditions that demand that social conditions are such that the agent has adequate 'live' options available, and the authority to pursue some other than that she in fact takes).
- 6) A necessary condition for autonomous agency is the constitutively relational condition (which demands that social conditions afford the agent *de facto* power).

However, it is not clear that we should accept the first part (1-3, above) of the argument. Recall the intuition pertaining to the control that autonomous agents have, and which informs premise 1: "Autonomous beings are in actual control of their own choice, actions and goals."⁴² This is plausible, and uncontroversial as Oshana claims. But we will see that the way she cashes this intuition out in terms of de facto power is not uncontroversial. There are two ways to understand the notion of control. The first notion of control is at work in Oshana's argument. The second notion is weaker. If this notion of control figures in Oshana's argument, then she cannot secure either the claim that the agents in her examples are non-autonomous, nor that the conditions for autonomous agency are constitutively relational. Oshana fails to provide good reason not to endorse only this weaker sense of control.

5.2.4 Two kinds of control

Oshana's concern is that the deferential wife Harriet and the monk lack a certain kind of control. Now, it is very plausible to hold that agents who occupy deferential or subservient roles do have a kind of control over their choices and actions, on some – perhaps many – occasions. They have the kind of control that compulsives lack, or that

⁴² Oshana, M. (2006) *Op.Cit.* p.3.

we lack if we have a bodily spasm, or suffer alien arm syndrome.⁴³ Oshana appropriates (somewhat liberally, as she acknowledges) from Fischer and Ravizza the terminology of guidance control, and the contrasting notion of regulative control.⁴⁴ The former kind of control, that deferential individuals and the monk plausibly possess, involves a kind of guidance. The agents' choices are causally connected, in the appropriate way, to their actions, and their actions to the events in the world, in a way that makes it appropriate to say that the agent was in control of her action.

Having guidance control means having the ability to govern one's choices and actions, on the basis of one's motivational states, such that one is 'in the driving seat' with respect to one's action. That is, one's mental states and bodily movements are not bypassed in the production of action, but are related appropriately to it. This kind of control is possible even if an agent cannot do otherwise than the action she in fact does.

Regulative control involves not merely that the agent can guide her choices and actions, but that she could guide her choices and actions other than as she in fact does; that she has the power not only to bring about the state of affairs she in fact does, but that she has the power to bring about an alternate state of affairs. Regulative control, then, requires that the agent has other 'live' options other than the option the agent in fact takes (in contrast to guidance control, which does not). So we have a distinction between two kinds of control:

⁴³ This is a syndrome in which the limb of the sufferer is experienced as behaving 'on its own' as if an alien body; 'its' actions are one's out of the agent's control, and the agent is alienated from it. See Lewis T.M.; McClain M.; Pittenger A. (1997) 'Alien Hand Syndrome and Sensory Ataxia: A Case Study of a Unique Presentation of the Disorder', *Archives of Clinical Neuropsychology*, 12 (4), pp.357-358.

⁴⁴ See Fischer, J. & Ravizza, M. (1998) *Op.Cit.* Fischer and Ravizza talk of guidance control being possessed by a 'mechanism' for choice and action – deliberation, say. This detail need not concern us here, and we can talk more generally of the kind of control the agent has, rather than her 'action producing mechanism'.

Moreover, Fischer and Ravizza talk about 'ability to do otherwise' in the *metaphysical sense* – that is, whether it is metaphysically possible, given the state of the world and the laws for the agent to do otherwise. Oshana is not concerned with this sense, but is most plausibly concerned with whether the agent could *rationally* or *prudentially* do otherwise, given her social context and the relations in which she stands to others; that is, whether the options are 'live' ones. For instance, she is concerned with whether, if the agent chose to do otherwise, she would face significant sanction or economic hardship. For instance, she demands that an individual should be able to pursue goals and values "different from those who have influence and authority over her, without risk of reprisal sufficient to deter her in this pursuit" (1998, *Op.Cit.* p.87). Of course, the agent could choose otherwise, and risk reprisal – but that there is a risk of reprisal means doing so may not be prudent; the option is not a 'live' option insofar as there remains such a risk. We should see Oshana as concerned with a social analogue to this metaphysical notion.

Guidance control: having the ability to guide one's choices and actions, on the basis of one's motivational states, such that these states appropriately cause the agent's choice and action. The agent is 'in the driving seat' with respect to her action.

Regulative control: having guidance control over what one does, and having the ability to exercise guidance control over an action other than the one the agent in fact performs (that is, the ability to do otherwise).

The following analogy helps to bring out the contrast:

Sally's car is a 'driver instruction' automobile with dual controls. We can imagine that the instructor is quite happy to allow Sally to steer the car to the right, but that if Sally had shown any inclination to cause the car to go in some other direction the instructor would have intervened and caused the car to go to the right.⁴⁵

Although Sally could not do otherwise than turn right, then, there is still a sense in which she is in control of turning right – this is the sense involved in guidance control. In contrast "Regulative control involves a dual power: for example, the power freely to do some act A and the power freely to do something else instead" (for example, the power to turn left and the power to turn right).⁴⁶

Of these two kinds of control, which does Oshana have in mind in the control intuition? Having considered guidance control, Oshana writes that "personal autonomy requires control of a more vigorous variety. To claim autonomy is to claim that a person has the power to determine how she shall live"⁴⁷ Indeed, Oshana's complaint about the Harriet and the monk is that social conditions do not permit the agent the power to choose or do otherwise. Thus Oshana demands a kind of regulative control for autonomous agency.⁴⁸

Now, this kind of control is more demanding than guidance control. Whereas guidance control requires only that the agent has a certain relation to her action, regulative control demands this and that she could have chosen and done otherwise. Many philosophers hold that agents can be autonomous whilst possessing only guidance

⁴⁵ *Ibid*, p.32.

⁴⁶ *Ibid*, p. 31.

⁴⁷ Oshana, M. (2006) *Op.Cit.* p.83.

⁴⁸ Indeed, in her (2003) ('How much should we value autonomy?' in *Autonomy* Frankel Paul, E., Miller Jr., F.D., Paul, J. (eds.) Cambridge University Press, pp.99-126) Oshana makes explicit that guidance control is not sufficient for autonomy, but that rather *regulative* control is required. See p.101.

control.⁴⁹ Understood in this way, then, the control intuition is not uncontroversial and is not accepted by laypersons or philosophers alike as Oshana contends. If the claim that regulative control is required for autonomy is to be accepted, we need to see that it is well motivated.

5.2.5. Is regulative control required?

However, this aspect of Oshana's argument is supported by little argumentation. There appear to be two strands of thought at work in motivating her commitment to regulative control. The first thought is that the alternative is inadequate. Oshana writes that "when we say a person is self-governing because she is in control of her actions and choices, we are saying more than that the person's actions coincide with the preferences and values that are her own. We are saying that the person has the power to determine how she shall live".⁵⁰

Now, it is right to maintain that self-governance and control amount to more than this. But to demand that control consists in more than the mere coincidence is not yet to demand that it must be regulative control: Oshana here presents us with a false dilemma. We can maintain that control at a minimum should consist in non-accidental coincidence of the agent's action with her preferences and values. Or, we may demand that there should be a reliable, non-deviant, causal link between the agent's preferences and values, and the action that the agent performs.⁵¹ Both of these claims amount to more than mere coincidence of the agent's preferences and action. If the agent is coerced into performing an action that happens to be consistent with – perhaps even

⁴⁹ For instance, compatibilists maintain that an agent can be autonomous whilst she cannot do otherwise. One might maintain that compatibilists are primarily concerned with metaphysical freedom – whether the world is such as to permit alternate possibilities. Oshana on the other hand is concerned with whether autonomy requires that social conditions are such as to permit alternative actions. It is possible that one might hold that the metaphysical ability to do otherwise (regulative control) is not required, but the agent must believe that social conditions are such that she could do otherwise. In response, we can say that insofar as she could not do otherwise, she lacks the de facto power to do otherwise, and hence can have only guidance control. It may be that social conditions are such as to lead the agent to believe that she cannot do otherwise. But this may thwart autonomy only insofar as the agent then fails to meet some of the content-neutral conditions: lack of endorsement or identification with her action, say. See Frankfurt, (1969) 'Alternate Possibilities and Moral Responsibility', *The Journal of Philosophy*, 66 (23) pp.829-839 for such a view.

⁵⁰ Oshana, M. (1998) *Op. Cit* p.82.

⁵¹ See Davidson, D. ((1980), *Essays on Actions and Events*, Oxford University Press p. 78) for an example of deviant causal chains in bringing about actions. In his example a climber contemplates letting go of a rope and ditching his companion. His nervousness at such a thought causes his grip to loosen and the rope to drop. The causal chain from his thoughts about letting go and his loosening his grip is a deviant one.

promotes – the values that she holds, she may not have the kind of control necessary for autonomy (namely, if it is not her endorsed motivational states that bring about the action).⁵² Or, recall the example from chapter 1, in which Gina accidentally knocks cola into her basket, in accordance with her grocery-shopping plans and choices. Gina’s action coincides with preferences that are her own. But this coincidence is accidental. When she is self-governing, we are saying more than that her actions coincide with her preferences and values. We can say that there is a non-accidental causal relation between the agent’s preferences and values, and her decision and action.

This first thought, then, does not support the claim that regulative control is required for autonomy. Rather, it supports the claim that self-governance requires more than accidental coincidence of preferences with action. Guidance control can satisfy this requirement; when an agent has guidance control, as we have seen, her motivational states appropriately cause the agent’s choices and actions.

The second thought that Oshana relies upon in support of her commitment to regulative control is the thought that characters such as Harriet and the monk are non-autonomous. If we grant this is so, it can be explained only by the claim that they lack regulative control – due to the social conditions in which they stand. But these intuitions themselves are controversial, as I will show in more detail. Unless the control intuition, or the claim that the agents in the case studies are non-autonomous, is supported by independent argument, her account suffers from being question-begging in a problematic way. This is made clearer if we consider the following expanded reconstruction of the argument:

Argument 1’

- 1) autonomous agents have control over their choices, will, and action such that they have the ability to do otherwise than they in fact do (the first control intuition, modified to regulative control)
- 2) certain roles (e.g. a deferential role) are incompatible with the agent having control over her choices, will and action such that she has the ability to do otherwise than she in fact does.
- 3) such roles (e.g. deferential roles) are incompatible with autonomous agency.

⁵² Although in discussing such an example, Frankfurt, H. (1969) *Op.Cit.* notes that it will sometimes be difficult – even impossible – to tell whether the agent is being motivated by her own (in the relevant sense) desires and preferences, or the presence of the coercion.

- 4) these roles are compatible with the internalist psychological conditions being met.
- 5) any failure of autonomy must therefore be because the required relational or ‘externalist’ conditions (that is, the conditions that demand that social conditions are such that the agent has adequate ‘live’ options available, and the authority to pursue some other than that she in fact takes) that supply the agent with ability to do otherwise fail to obtain.
- 6) A necessary condition for autonomous agency is the constitutively relational demand for social conditions that supply the agent with ability to do otherwise.

The conclusion reached states that autonomous agency requires that social conditions supply the agent with ability to do otherwise (which Oshana sometimes puts in terms of the agent having de facto power). The first premise, initially, stated:

- 1) autonomous agents have control over their choices, will, and action (the control intuition)

When we expanded this premise so that it accords with the sense of control Oshana has in mind, and it reads:

- 1) autonomous agents have control over their choices, will, and action such that they have the ability to do otherwise than they in fact do (the control intuition, modified to regulative control)

Thus, when expanded, we see that the conclusion reached in 6 is in fact assumed by the first premise, with the modified control intuition. The conclusion that autonomy requires that the agent has the ability to do otherwise (as permitted by certain social relations) only follows given the assumption (in premise 1), that the control required for autonomy involves the ability to do otherwise. If Oshana is concerned with regulative control, as she appears to be, the argument for her constitutively relational condition is question-begging.

What happens if we start with the weaker sense of control? Can the relational condition be secured starting from a premise that assumes only guidance control as

necessary for autonomous agency? If this is the kind of control at issue, then the argument for the constitutively relational conditions looks like this:

Argument 1''

- 1) autonomous agents have guidance control over their choices, will, and action (the control intuition, modified to guidance control)
- 2) certain roles (e.g. a deferential role) are incompatible with the agent having guidance control over her choices, will and
- 3) such roles (e.g. deferential roles) are incompatible with autonomous agency.
- 4) these roles are compatible with the internalist conditions being met.
- 5) any failure of autonomy must therefore be because the required relational or 'externalist' conditions (that is, the conditions that demand that social conditions are such that the agent has adequate 'live' options available, and the authority to pursue some other than that she in fact takes) that supply the agent with ability to do otherwise fail to obtain
- 6) A necessary condition for autonomous agency is the constitutively relational demand for social conditions that permit ability to do otherwise.

However, if we start with a premise that assumes only guidance control is required for autonomy, then Oshana cannot establish that the constitutively relational condition (RelAgency2) is necessary for autonomous agency. First, the argument is unsound: premise 2 of this argument is false – occupying a role of subservience is not incompatible with having guidance control. Second, the premises (5 and 6) that make reference to the relational, externalist conditions that permit ability to do otherwise do not follow from earlier premises – it does not follow from the (false) claim that guidance control is incompatible with certain roles that this is because of the failure of certain social conditions or relations that supply ability to do otherwise. Given that guidance control does not require the ability to do otherwise, it will not follow from a (false) premise about the incompatibility of certain roles with this kind of control that a necessary condition for control is the ability to do otherwise.

Might Oshana escape this line of argument by insisting that her argument need not start from the control intuitions? Rather, we can start from the intuition that the monk and the deferential individual, Harriet, are not autonomous – such roles are incompatible with autonomous agency. Because agents who occupy such roles nonetheless meet the internalist conditions, their non-autonomy must be explained by

the failure of certain social conditions to obtain: the absence of feasible options removes the agent's ability to freely choose and manage her life. If the argument starts from the intuitions about the agents, then she can avoid the charge of question-begging.

However, running the argument in this way means that a great deal of weight is placed on the intuitions about the agents in question. As noted, there are many who are unswayed by the claim that such agents are non-autonomous. For example, Andrea Westlund writes that

While I share Oshana's conviction that many relations of subordination (including this one) are substantively criticisable on feminist and other grounds ... if (*ex hypothesi*) a fundamentalist woman *does* freely and authentically accept a condition of social and personal subordination, it seems equally problematic to assume that her condition as subordinate, in and of itself, undermines her status as self-governing agent. It may be that standard internalist views leave something to be desired in their handling of such cases. But if we want to construct the most formidable test case for an internalist view, we need to be more attentive to possible differences [in degrees of autonomy] between self-subordinating characters.⁵³

Likewise, Benson has suggested that groups of individuals who act in accordance with oppressive norms might be internally diverse, with some agents who so choose showing a far greater engagement with their reasons for action than others.⁵⁴ His point carries over; there may be significant differences between individuals who choose to occupy certain social positions of subordination, and some may strike us as more autonomous than others. Gerald Dworkin is another prominent theorist who denies the intuition that agents who occupy roles of subordination are therefore non-autonomous.⁵⁵

Given the contention surrounding the claim that agents in such roles are autonomous, to start the argument from such intuitions is unsatisfactory. Of course, Oshana could attempt to substantiate the intuition. But if she does so by appeal to the control intuition, then as we have seen, her argument will be question-begging. Perhaps an alternative way of substantiating the intuition is to appeal to the lack of options such agents are presented with. For instance, Oshana might bolster the intuition with a claim that:

⁵³ Westlund, A. (forthcoming) 'Rethinking Relational Autonomy', p. 5.

⁵⁴ See Benson, P. (2005b) *Op. Cit.*

⁵⁵ See Dworkin, G. (1988) *Op. Cit.*

the self-governing individual must have access to an adequate assortment of options. ... Because self-governance is governance over matters of central importance to human life, the options available must be relevant to the development of a person's life and they must be ones a person can genuinely hope to achieve.⁵⁶

Now, the claim that an agent must have an adequate range of options cannot be motivated by the claim that this is required to ensure the agent regulative control, in the sense with which she is concerned – we will return, with another iteration, to the question-begging concern. But the claim that an adequate range of options are required need not only be justified in this way; it has independent intuitive plausibility.

However, if Oshana's argument moves in this direction, it becomes much less clear that she is able to maintain her rejection of the 'apples and oranges' objection. Those who offer accounts of autonomy that appeal to the psychological states of the agent (internalist accounts) are concerned with the range of options the agent has only insofar as this prevents the agent from meeting the internalist conditions. In making the range of options necessary for autonomy independently from the impact of the range on the internalist conditions, Oshana tends towards an understanding of autonomy; one concerned less with self-governance, and more with a particular, valuable way of conducting one's life. Indeed, of this sense of autonomy Oshana notes that "not everyone will include an autonomous life [namely, one that includes the social roles and conditions she specifies] among the goals that she regards as integral to well-being".⁵⁷ With this claim, we see that she views autonomy as a valuable way of life, rather than as a capacity that entitles agents to some important normative benefits.

In this, Oshana departs from a conception of autonomy which plays a role in grounding respect, identifying candidates for political participation and protection from interference: the notion of autonomy that plays these roles extends to all agents, even if they adopt conceptions of the good or social positions other than those which ensure a range of options, and other than those consonant with the value of substantive independence. It is plausible that a range of options will be involved in a flourishing life, and one legitimate use of the term 'autonomy' is to such lives. But this is not the way in which the internalist theorists are using the term: they are talking apples, Oshana oranges.

⁵⁶ Oshana, M. (2006) *Op. Cit*, pp.84-85.

⁵⁷ *Ibid*, p.104.

To summarise: unless Oshana can establish with independent argument that regulative control is required for autonomy, then the claim that Harriet and the monk are non-autonomous need not be accepted. The claim that social conditions need to be such as to secure a range of options over which the agent has *de facto* power cannot be substantiated. The intuition that the agents themselves are non-autonomous is insufficiently strong to provide an independent starting point for the argument. And if Oshana substantiates the intuition with reference to the value of having a range of options (which is, plausibly, valuable) then she has made a significant departure from the sense of autonomy at work in the psychological, internalist accounts; the ‘apples and oranges’ objection returns. There is no good reason, then, to accept the first two clauses of the relational condition (RelAgency2). What of the claim pertaining to the agent’s power counterfactually, such that she has the authority to maintain control in the possible event that others attempt to intervene?

The following considerations count against this condition. First, recall that the authority intuition is connected to the control intuition in the following way: being subject to the potential intervention of others undermines autonomy because it prevents the agent from having the range of ‘live’ options that the autonomy requires. If only guidance control is required for autonomy, and other options are not required, then insofar being under the threat of potential intervention (‘domination’ by others) undermines autonomy it does not do so by removing the agent’s *de facto* power, understood as the ability to freely act otherwise. If the authority intuition is relevant to autonomy (which I think it is), it is not relevant in the way that Oshana sets out; namely, it is not relevant because of its connection to the range or feasibility of options that the agent has.

A second consideration arises, which will be addressed in more detail in the next chapter. Briefly put, it is plausible to hold that in cases in which interventions of others are problematic, one of the claims we should wish to make is that the intervention is problematic precisely because it overrides the agent’s autonomy. Arguably, agents retain their status as autonomous beings when subject to intervention, and therefore when subject to the possibility of intervention; it is rather that this status is not respected (and subsequently their liberty, well-being, interests, and so on are thwarted). It is (in part) because agents are autonomous that some interventions, or threatened interventions, are problematic. If such interventions or potential interventions remove

autonomy, such claims cannot be made. This thought will be returned to in greater detail in chapter 6.

Finally, it is worth noting that once again, a weaker understanding of the authority intuition is available; namely, one which focuses on the kind of ‘intrapersonal authorising’ of an agent’s deliberation, choice and action. Autonomous persons have authoritative control of their choices and actions, on this understanding, when they see their choices and actions as ‘their own’, and issuing from the deliberative processes that they endorse. This is the kind of authority that Benson has in mind when he writes that autonomous agents have the ‘authority to construct, stand by, and speak for their reasons for acting’.⁵⁸ This weaker notion of authority, which focuses on the agent’s sense of accountability for her actions, may well be necessary for autonomy; but it does not demand that the agent stand in certain social relations.

Neither the authority nor the control intuitions are uncontroversial, then. Oshana has given us insufficient grounds to endorse the strong readings of these intuitions that inform her constitutively relational condition (RelAgency2).

5.3 Substantive concerns

I have argued that the considerations that motivate Oshana’s constitutively relational condition need not be heeded: first, the psychological ‘internalist’ accounts do not paint a picture of autonomy as isolated from the impact of social relations. We saw as much in looking at Benson’s account. Second, the claim that autonomy requires social relations that provide a range of live options and de facto power (namely, ability to do otherwise) can only be secured insofar as the requirement for regulative control is motivated. But it is not. Moreover, to independently motivate the intuitions without recourse to regulative control – and rather to the value of a range of options – is to return to face the ‘apples and oranges’ objection. Third, the weaker kind of control – guidance control – involves more than the accidental coincidence of preference and choice, as Oshana demands. Finally, there is reason to doubt that the authority intuition that informs her relational condition need be accepted in its strong form.

This removes much of the motivation for positing constitutively relational conditions. In this section, I briefly return to the concern about values in an account of

⁵⁸ Benson, P. (2005b) *Op.Cit*, p.133.

autonomy that I raised at the beginning of Chapter 4, and the relation between constitutively relational conditions and value-laden content.

We have seen Christman's concern that incorporating values into an account of autonomy has troubling consequences, in terms of the exclusion of those who fail to subscribe to the stipulated values. He has argued that Oshana's account in particular faces this objection. Now, Christman thinks that this problem is one that will face any account that attempts to incorporate constitutively relational conditions for autonomy. It is one thing, he says, to hold that social relations are causally necessary for autonomy, and thus

to claim that social conditions that enable us to develop and maintain the powers of choice ... are part of the background requirements for the development of autonomy....⁵⁹

However, he claims that we should reject constitutively relational conditions for autonomy, because to demand certain social relations means also demanding of agents that they commit to a certain value perspective (namely, one that sees those social relations as valuable) in order to meet the standards for autonomy. The move is ultimately problematic, Christman claims, because

viewing [social] relations as constitutive of autonomy implies that certain values ... are valid for individuals even if they authentically and freely reject them.⁶⁰

That is to say – it is troubling to require certain social relations for autonomy, because this has the consequence that those who have, on the basis of competent deliberation, rejected the set of values that those social relations enshrine can nonetheless be evaluated and excluded from the class of autonomous agents according to whether their lives cohere with the values that they have rejected. For example, Christman maintains that Oshana's constitutively relational account imports value commitments:

views like Oshana's rest upon the claim that certain value commitments – such as the view that I must obey my superiors unconditionally – are conceptually inconsistent with autonomy.⁶¹

⁵⁹ Christman, J. (2004) *Op.Cit.* p.152.

⁶⁰ *Ibid*, p. 152.

⁶¹ Christman, J. (2006) *Op.Cit.* p.151.

Such a consequence, he claims, is not compatible with the liberal commitment to there being a plural set of inconsistent but reasonable values according to which one might live. Rather, her account privileges the value of substantive independence.

It is further problematic, because maintaining that only those agents who subscribe to certain values are autonomous risks excluding from the range of normative benefits that attach to autonomy many individuals who are otherwise competent and capable of self-governance. For instance, agents who fail to stand in social relations that permit them the ability to do otherwise (having chosen this or not) will not, on such a view, possess

the status marker of an independent citizen whose perspective and value orientation gets a hearing in the democratic processes that constitute legitimate social policy.⁶²

Such agents will not be candidates for political participation. Nor will they be protected from paternalistic interventions, for should an agent fail to occupy the social relations that Oshana demands, then:

her lack of relational autonomy – should we accept that view – would allow other agents and representatives of coercive social situations to intervene to relieve her of this burden and to restore her autonomy (at least in principle).⁶³

Finally, Christman mentions the role of autonomy in grounding respect, and Friedman clearly articulates the concern that value-laden views may pose problems for this dimension:

An account of autonomy that is too demanding will prompt persons to regard a greater number of others as failures at personhood and thereby reduce the number of others they regard as respect worthy.⁶⁴

Constitutively relational conditions for autonomy, in short, do not appear to be well-placed to play the normative roles that account for autonomy's value. At least – this will be so if it is correct that constitutively relational conditions are committed to the claim

⁶² *Ibid*, p.157.

⁶³ *Ibid*, p.157.

⁶⁴ Friedman, M. (2003) *Op.Cit.*, p.23.

that an agent is autonomous only insofar as she chooses and lives in accordance with some value.

So, insofar as Oshana's view incorporates the value of independence, judging the autonomy of agents according to the degree to which their lives instantiate this value, she incorporates into her view "an ideal that those who choose strict obedience or hierarchical power structures have decided to reject".⁶⁵ Those who reject this idea may not secure the normative benefits, as set out above. Note that in addition to these problems of exclusion, such views require

that the content of the judgments [or choices] in question reflect a high regard for personal independence or other values of the sort that views such as Oshana's demand of the autonomous agent.⁶⁶

If constitutively relational views end up placing constraints upon the contents of what can be chosen autonomously (rather than on the social relations which those who qualify as autonomous agents must stand), we will return to the kind of concerns, considered in Chapter 4, facing those views which place substantive constraints on the contents of autonomous choice; namely, the problem of finding a satisfactory way of constraining choice without ruling out a significant portion of those agents deemed to be un-controversially autonomous (what I called the problem of false negatives).

5.3.1 The scope of Christman's objection

I have set out Christman's claim that we should not accept any account that offers constitutively relational conditions, insofar as they incorporate values or ideals.

His claim is that we should reject constitutively relational conditions for autonomy:

It is ... an ultimately problematic move ... to claim that being autonomous means standing in proper social relations to surrounding others and within social practices.

Now, an important point to note here is that Christman is not merely rejecting value-laden conditions; he also claims that we should not accept any constitutively relational conditions for autonomy. This is because Christman conflates value-laden conditions

⁶⁵ Christman, J. (2004) *Op.Cit*, p.151.

⁶⁶ *Ibid*, p.151.

with constitutively relational conditions. Elsewhere he characterises relational conceptions of autonomous agency as those

according to which persons are autonomous only when they have particular kinds of value commitments.⁶⁷

But insofar as it is possible to give relational conditions for autonomous agency that do not incorporate any values, Christman's concern can be avoided. And it seems at least possible that one could offer value-neutral constitutive conditions. For example, one might hold that an agent must stand in social relations in which she is a party to the exchange of reasons for action – a demand possible from the perspective of any value commitment.⁶⁸ Such a view need not 'smuggle in' any values nor demand that the agent has chosen any particular value commitment or way of life. Perhaps these constitutively relational conditions are not, in fact, ones that we should accept for autonomy. But insofar as they are not value-laden, any objections to them will not be grounded in concerns such as Christman's.⁶⁹

In the next chapter, however, I will argue that there are other reasons for rejecting constitutively relational conditions for autonomous agency. So even if value-neutral conditions for autonomous agency can be given, we should not accept them. But I will go on to argue, in Chapter 7, that there are constitutively relational conditions for autonomous action which do not incorporate any substantive value commitments. Christman's conflation of relational conditions with value-laden commitments, then, is a mistake.

5.4 Summary

In this chapter, I started by considering whether Benson's view might provide us with constitutively relational conditions for autonomous agency. I argued that it should not be so understood, but that important lessons can be learnt from his account nonetheless: namely, that value can be incorporated without directly constraining choice; and that

⁶⁷ Christman, J. (2006) Book Review: Marilyn Friedman, *Autonomy Gender, Politics* (2003), in *The Journal of Value Inquiry*, 40, pp.137-143, see p.139.

⁶⁸ Westlund, A. (2003) *Op.Cit* argues for such a condition.

⁶⁹ Westlund, A. ('Rethinking Relational Autonomy', forthcoming) argues that constitutively relational accounts need not be value-laden, using her own view as an example. However, her view of autonomy – which gives central place to an agent's disposition to engage with critical perspectives – appears to me to be merely causally rather than constitutively, relational, in a way similar to Benson's.

views which focus on the psychological states of the agent can take into account the impact of social relations upon autonomy. This removed some of the motivation for Oshana's constitutively relational account, which I then considered. I argued that her view was further lacking in motivation: the claim that regulative control is required for autonomous agency remained unargued for. Without any independent support for this claim, the argument for her constitutively relational condition is question-begging. Finally, I set out the concern from Christman that constitutively relational conditions smuggle in values. Insofar as this is so, they are unable to play a role in the normative frameworks that in part explain its value.

I suggested that this conflation is mistaken: it is at least possible that one might offer constitutively relational conditions without incorporating value-commitments. Christman's concern will not address such views. In the next chapter, I offer an objection that faces all relational or 'externalist' accounts of autonomous agency, whether or not they incorporate substantive value commitments.

Chapter 6. An argument against constitutively relational conditions for autonomous agency

In the last chapter, I looked at two attempts to take into account the way that social conditions might be constitutive of the conditions for being an autonomous agent. The first attempt, by Benson, most plausibly amounted to the claim that social relations might have a causal impact upon the sense of competence of autonomous agents. The second view, from Oshana, involved the claim that agents cannot be autonomous unless they occupy certain social conditions; those which enable a range of options and secure certain powers for the agent. Having considered specific problems with each of these views, in this chapter, I will offer a general argument against constitutively relational conditions for autonomous agency, namely, any condition that has the form:

(RelAgents) A necessary condition for being an autonomous agent is that the agent stands in social relations S

Focusing on the role that autonomy plays in the normative framework of protecting from paternalism, I will argue that no account which incorporates a constitutively relational condition can play this normative role. This will motivate the turn, in the next part of the thesis, to consider the relational conditions for autonomous action.

6.1 Paternalism and autonomy

An account of autonomy, recall, should be able to play important normative roles: I have mentioned three of these, namely, the roles of grounding respect, identifying the candidates for political participation, and those who should be protected from paternalistic interventions. In this chapter I will focus on the latter framework: that in which the notion of autonomy is supposed to play the role of marking out those agents who should be protected from a certain kind of intervention.

One of the intuitions that Oshana is aiming to capture, with her relational account, is that “autonomy is the good that paternalism fails to respect”.¹ Likewise,

¹ Oshana, M. (1998) *Op.Cit.* p.82.

Christman holds that “the idea of autonomy set[s] the boundaries of ... anti-paternalism in principles of justice”.² He goes so far as to claim that “all agree ... that another function of the concept of autonomy is to mark out the parameters within which a person is [or should be] immune from paternalistic intervention”.³

Assuming for now that this is correct (I will later say more to substantiate this assumption), I want to look in more detail here at paternalism, and its relation to autonomous agency – namely, the features in virtue of which autonomous agents should be protected from paternalistic interventions. My aim is to show that an account which incorporates a relational condition such as (RelAgents) makes reference to the wrong kind of features to play this role. Thus insofar as we want the class of autonomous agents to be coextensive with the class of agents entitled to protection from paternalism, we should not accept a constitutively relational condition for autonomous agency.

6.1.2 Paternalistic interventions

When is an intervention a paternalistic one? We already have part of the answer from Oshana: an intervention into a choice or action is paternalistic when the intervention fails to respect the autonomy of the agent. But we want to know more about how such an intervention fails to respect the agent’s autonomy, and how this kind of intervention differs from other cases in which autonomy is not respected.

a. Paternalism vs. other wrongdoing

Lying, deceiving, stealing are all ways of interfering with another’s action, and arguably constitute failures of respect for autonomy. The Kantian diagnosis of the problem with such interferences is that, in (say) lying to another, that agent is treated not as an end, but as a mere means. Insofar as the lie is intended to be part of a manipulation of the agent for the ends of the liar, the agent is used for the furthering of the ends of the liar. Moreover, the agent cannot consent to being used in such a way – in virtue of being lied to the agent does not know, so cannot share, the ends of the liar. Likewise, if an agent is stolen from, she is not respected as an end – her goals and aims – in particular, those with regard to her property – are not respected, but disregarded.⁴ These kinds of

² Christman, J. (2004) *Op.Cit.* p.147.

³ *Ibid*, p.157.

⁴ See e.g. Korsgaard, C. (1996b) *Op.Cit* for such claims. But see Buss, S. (2005) *Op.Cit* and Parfit, D. ((2002) ‘What we Could Rationally Will: I. Rational Consent’, *The Tanner Lectures on Human Value*. pp.287-369, <http://www.tannerlectures.utah.edu/lectures/atoz.html#p> esp. at p.

autonomy-violation, then, are clear ways of failing to respect the autonomy of an agent; such wrong-doing disregards the agent's own plans and determinations.

Paternalistic interventions can be distinguished from these kinds of autonomy-violations, because paternalistic interventions are motivated by the good of the agent – they are intended to further the interests of the agent. As Dworkin puts it, paternalistic interventions are those which involve “interference justified by reasons referring exclusively to the welfare, good, happiness, needs, interests or values of the person being coerced”.⁵ In one sense, then, the interventions are benevolent; they appeal to the interests or good of the agent who is subject to the coercion or the intervention. In what way, then, should we understand such interventions as violations of autonomy?

b. The wrong of paternalism

Whilst the interests of the agent are made reference to in cases of paternalistic intervention, there is nonetheless a violation of autonomy. This is because, in such interventions, one agent “substitutes their judgment for another person's. There is an usurpation of decision making”.⁶ The agent's own judgment about what will serve her interests, or further her good, or about what should constitute her ends or goals, is replaced (for the purposes of determining what course of events comes about) by the judgment of another (the paternalist). The intervention of the other, then, is in this sense against the will of the agent, even though it is motivated by concern for her interests. Thus the autonomy of the agent, in determining her choices and actions, is undermined. Her choices are not respected, and her decisions about the actions that she seeks to bring about are overridden.

297) for complications concerning the matter of autonomy-violation and treating as a mere means.

⁵ Dworkin, G. (1972) *Op.Cit* at p. 65. Note that Dworkin characterises paternalistic interventions as those that are exclusively motivated by reasons pertaining to the good, or welfare or interests of the agent. This understanding is narrower than that used by Joel Feinberg, who maintains that an intervention is paternalistic so long as one of the reasons for intervention makes reference to the welfare of the agent (this may not be the decisive reason, and there may be other reasons): “it is always a good and relevant (though not necessarily decisive) reason in support of a criminal prohibition that it will prevent harm (physical, psychological or economic) to the actor himself” (Feinberg, J. (1986) *Harm to Self: The Moral Limits of the Criminal Law, Volume Three*, Oxford, Oxford University Press, p.4).

For another formulation of paternalistic intervention that is closer to Mill's original injunction against paternalism, see Richard Arneson, (1980) ‘Mill Versus Paternalism’, *Ethics*, 90 (4), pp.470-48. I here work with Dworkin's narrower conception, the reasons for this strategy will be apparent in the following. A useful discussion of the different understandings can be found in Clare Chambers' (2008) *Op. Cit.* Chapter 6.

⁶ Dworkin, G. (1983) *Op.Cit.* p.107.

Some of those who argue that paternalism can sometimes be justified consider whether the benefits of overriding an agent's choice can be outweighed by the problems of autonomy-violation. It is worth noting that Mill claimed that the balance of costs and benefits would (almost) never fall on the side of paternalism – this thought in part is motivated by the claim that the agent knows her own interests, and what will serve them, better than any other.⁷ This claim seems to be falsified on a not irregular basis; we are often poor judges of what will serve our interests best, perhaps as a result of misinformation, or wishful thinking and other forms of motivated irrationality, or failures of self-knowledge. One might hold, however, that the moral harm of paternalism is sufficiently great that it is never justified – or justified only rarely; such as when severe harm will befall the agent, or when the agent's life is at stake. Let us characterise a paternalistic intervention as follows:

an intervention in the choice and action of an autonomous agent, exclusively for the sake of the agent's good, which does not treat the agent as if she is capable of making her own decisions.

It is important to note that there are two ways of understanding the reference here to intervention which involves failing to treat the agent 'as if she were capable of making her own decision'. On one reading, to so treat an agent may be to treat her as if she lacks the capacity to make a competent decision; she is treated as one might treat a small child or a patient in a coma. On another reading, to so treat an agent may be to treat her as lacking the epistemic state – in terms of the beliefs she is possessed of – to make her own decision. An example of the first: Myrna recognises that Ignatius has made a decision of his own, but attributes little value or significance to his exercise of agency. She overrides Ignatius' decision as she would that of a young child. An example of the second case: Dorian sees that Timmy doesn't realise that the bridge he is about to cross will collapse and that he will henceforth plummet to his death; Dorian stops him from crossing, so that he can inform him of the circumstances. (An alternative possibility is that the agent is recognised to have the relevant capacities, but is temporarily unable to

⁷ See Mill, J.S. *Op. Cit.* "with respect to his own feelings and circumstances, the most ordinary man or woman has means of knowledge immeasurably surpassing those that can be possessed by anyone else", at p.207. It is also important to note that Mill's opposition to paternalism extended to the claim that the only case in which intervention with an individual's action is justified is in cases in which harm will come to others. The permission for such interventions is enshrined in his harm principle: "the only purpose for which power can rightfully be exercised over any member of a civilized community, against his will, is to prevent harm to others" p.135. Note, finally, that I say 'almost never', because Mill held that on occasions on which an agent chooses to forsake his freedom (for example, by selling himself into slavery), paternalistic interventions may be justified.

exercise them, due to (for example) drunkenness. I do not explore this case in detail here: it is plausible that such a case can be understood as one of ‘soft’ paternalism, as I set out below).

Now, in line with these two ways of understanding the sense in which one may treat another as lacking competence to decide, Dworkin distinguishes between ‘hard’ and ‘soft’ paternalism.⁸ ‘Soft’ paternalism involves interventions or impositions in cases in which the agent acknowledges – or would acknowledge – her fallibility, or recognises the potential for minor cognitive deficiency on her part, such that she would acknowledge that the course of action imposed upon her or into which she is coerced is one that she would, deficiencies aside, assent to. ‘Hard’ paternalism, on the other hand, is the imposition of or intervention with a course of action, when it cannot be said of the agent interfered with that, with further reflection, or more acquaintance with the facts, she would consent to it. In cases of ‘hard’ paternalism, then, we have one agent imposing some course of action, informed by some conception of the good, on another agent who stoutly rejects that conception, and the action justified by it.

6.1.3 Aside: clearing up a dispute

Note that Dworkin and Christman appear to use the term ‘paternalism’ differently: Christman holds that all autonomous agents are – or should be – immune to paternalistic intervention. This suggests that his view is that all paternalistic interventions are unjustified.

On the other hand, Dworkin holds that, whilst “there must be a violation of a person’s autonomy for one to treat another paternalistically”,⁹ some interventions can nonetheless be justified. On this view, there will be some cases in which the good that the agent accrues as a result of the intervention is sufficiently great to justify the overriding of their autonomy. Thus there seems to be a disagreement about what paternalism in fact is.

However, clarification of the claims at stake helps us see that there is, in fact no substantive dispute: Dworkin’s considered position¹⁰ is that only soft paternalism can be justified – interventions can only be justified in cases in which the agent, under

⁸ See Dworkin, G. (1972, 1983), *Op.Cit.*

⁹ Dworkin, G. (1972) *Op.Cit.* p.123.

¹⁰ Dworkin, G. (1983) *Op.Cit.* p.112.

motivationally similar circumstances, would consent. Christman's conception of autonomy includes a counterfactual: an agent is autonomous when she meets (among other conditions) the condition demanding that (holding constant the agent's significant motivational states) were she to reflect on the acquisition of the motives upon which one acts, she would endorse it.¹¹

With these two considerations in mind, we can thus see the views of Christman and Dworkin aligning, despite the terminological differences: intervention cannot be justified in cases in which one agent, (A), imposes a course of action, as informed by a conception of the good upon another agent, (B), where (B) does not, cannot and would not accept that conception of the good, nor the actions that it endorses in the case at hand. In cases in which the agent would endorse the intervention, she is autonomous in acting in accordance with the intervention, on Christman's counterfactual view. In cases in which the agent would not endorse the intervention, she is not autonomous on Christman's view, so is paternalistically coerced. And the intervention is a 'hard paternalistic' one, so cannot be justified, for Dworkin. It is with these cases of 'hard paternalistic' intervention that I will be concerned in the following.

6.1.4 The 'core case'

Christman and Dworkin, amongst others, are agreed that there are certain cases in which intervention with the agent's action, for the sake of her own good, are not justified. These are cases which have the structure of 'the core case'. It will be useful to have this case set out:

Agent A and agent B are both reflectively competent agents. Agent B is subject to coercive intervention or imposition, exclusively for the sake of her own good. Were she to have further facts, or were she to reflect further she would not consent to this intervention. The action imposed upon her, by A, enshrines a conception of the good to which B does not, and would not, upon further reflection, subscribe.

This is a clear-cut case in which one agent is imposing her conception of what is good for another upon that agent. What is important for present purposes is to identify in virtue of what features, broadly, intervention in the core case is unjustified. It will be instructive to see if the relational conception is adequately placed to account for what is

¹¹ And Christman's view, like Frankfurt's involves a dispositional view of desires and motives: X desires p if X pursues (under certain circumstances) p. This is a weak sense of motive, pertaining purely to the motivational 'push', rather than any phenomenal sense of desire or urge.

problematic about the core case. If it cannot, then whatever else is said about the other, more disputed territory (that of ‘soft’ paternalism), relational conceptions of autonomous agency will be unsuitable for this theoretical purpose. There are two plausible claims that the relational theorist should be able to make:

- i) the intervention in the core case is unjustified¹²
- ii) the intervention in the core case is unjustified because it violates the agent’s autonomy.

Let us start by noting that in the core case, what is objectionable is that one agent is intervening with the choices and actions of another, although (as stipulated) the choices of each display the same ‘formal features’. That is to say, the agents both have certain capacities, such as those for rational belief formation and choice. The agents are deliberatively adept, and are not in a position in which further reflection or facts would lead to a revision of choice. The agent who is subject to intervention has made a decision that is as good – procedurally speaking – as the decision of the intervener. It is due to the substantive content of the decision that the intervention is deemed to be justified by the intervener. In so intervening, there is a usurpation of decision-making; A’s decision usurps B’s.

Now, I have already endorsed the claim that we should not accept an account according to which a consideration in determining the autonomy of a choice is whether or not the choice accords with some substantive value. Combined with this thought, we have the claim that in the core case it is in virtue of the procedural ‘formal features’ that each agent possesses that the intervention appears intuitively unjustified. (And, conversely, it is when these formal features are lacking – when the agent lacks deliberative competence or reliable belief forming capacities – that intervention is most likely to be justified.) Thus it is the agent’s competence at decision making – her capacity for self-governance in deliberation and choice – that is important in

¹² Danny Scoccia rejects this thought, arguing that there are some cases in which hard paternalism can be justified (namely, cases in which non-terminally ill sound of mind individuals choose to end their lives). See Scoccia, D. (2008) ‘In Defense of Hard Paternalism’ *Law and Philosophy* 27, pp.351-381. For those who think that hard paternalism can sometimes be justified, a weaker version of these claims can be utilised:

- i) there is pro tanto reason against intervention in the core case
- ii) there is pro tanto reason against intervention in the core case because the intervention violates the agent’s autonomy.

It should be clear in setting out my arguments that they apply to these weaker claims also.

determining whether an intervention for the agent's own good is unjustified, and paternalistic.

We saw earlier that theorists wish to claim that such paternalistic interventions are violations of autonomy. We can now see that this claim, in part, involves the claim that the intervention is unjustified because it does not treat the agent as capable of making her own decision. And the claim that the intervention is unjustified is informed by the thought that the substantive content of a decision is not sufficient to warrant the overriding of an agent's judgment. When an agent has decision making capacities, and exercises these in coming to a decision, this decision should not be replaced by the decision of another who also possesses the capacities and abilities to make reflective decisions on her own behalf.¹³

Now, it is important to note that the features that are informing both the claim that the intervention in the core case is unjustified, and that it is unjustified because it violates the agent's autonomy, are 'internalist' features. 'Internalist' features are those which pertain to 'what goes on in the agent's head', as it were; aspects of her psychology, or competence in practical deliberation.

Insofar as it is the notion of autonomy that is doing work in making sense of why the intervention is unjustified, then, it is an internalist conception of autonomy.¹⁴ That is, the agent's psychological competences – her ability to deliberate and choose – are what make the usurpation of her choice and violation of her autonomy by another agent, objectionable. This is the crucial point, and the point that poses problems for the relational conception of agency.

6.2 Relational conceptions and the core case: the problem of 'externality'

I want to now consider whether a conception of autonomous agency that incorporates a constitutively relational condition of the form:

(RelAgents) A necessary condition for being an autonomous agent is that the agent stands in social relations S

¹³ The kinds of values that autonomous choice embodies – such as symbolic choice, for example, as discussed in chapter 1 – may be bound up with the value of making one's own choice, and the disvalue of it being over-ridden.

¹⁴ The kinds of conditions that are offered by internalist conceptions are those discussed in chapter 2 – competence, conditions, structural and procedural conditions, for example.

can make sense of the two plausible claims about the core case:

- i) the intervention in the core case is unjustified
- ii) the intervention in the core case is unjustified because it violates the agent's autonomy.

The relational condition, recall, is intended to supplement the necessary procedural conditions, thus yielding jointly sufficient conditions for autonomy. It will be instructive to consider these claims with reference to a specific relational conception – thus I will consider what Oshana's conception of autonomy, which incorporates the relational condition (RelAgents2), below, can say with regards these two claims.

(RelAgents2) A necessary condition for being an autonomous agent is that the agent stand in social relations in which she has de facto power to choose and manage her life; in which her choices are from an adequate range of options; in which she has power counterfactually, such that she is able to maintain control over her life even in possible scenarios in which others attempt to intervene.¹⁵

On this view “autonomy calls for a measure of substantive independence from other persons and from social roles and traditions of a variety deemed to be inhospitable to autonomy”.¹⁶ Autonomous agents must stand in social relations that ensure this measure of substantive independence.

Let us consider, then, what Oshana's relational condition has to say about ‘the core case’, in which an agent, B, has her decision overridden by another agent, A, despite the two agents both possessing the relevant capacities and deliberative abilities. If this case is to be considered, as I believe it ought to be, as an instance of paternalism, and if we are to hold that paternalism violates an agent's autonomy, then we must be able to say of this agent that she is autonomous, and is not being treated as such, albeit for the sake of her own good. Can the constitutively relational conception of autonomous agency deliver the claims that i) this intervention is unjustified, and that ii) it is unjustified because it is a violation of autonomy?

¹⁵ These conditions are set out in Oshana, M. (2006) *Op Cit.* pp. 83-90. I discussed the details of her view at length in the previous chapter. It is worth noting here that Oshana's reference to intervention is to arbitrary interventions – so does not itself rule out paternalistic intervention.

¹⁶ Oshana, M. (2006) *Op Cit.* p.72.

I deal with ii) first. At first glance, it is simply not clear what an account that incorporates a condition such as (RelAgents2) can claim. The relational theorist will need more information about the social context in which the agent is operating, namely, whether the conditions present an adequate assortment of options, enable protection of rights and powers, and so on. Only then can she determine whether or not the agent is an autonomous agent, and only then could the claim be made about the intervention being a violation of autonomy (or not). For example, Oshana will want to know not simply about the agent's competence as a deliberator, but also about whether the agent stands in certain social conditions – those which ensure a certain range of realistic, 'live' options, and a level of protection from the interference of others.¹⁷

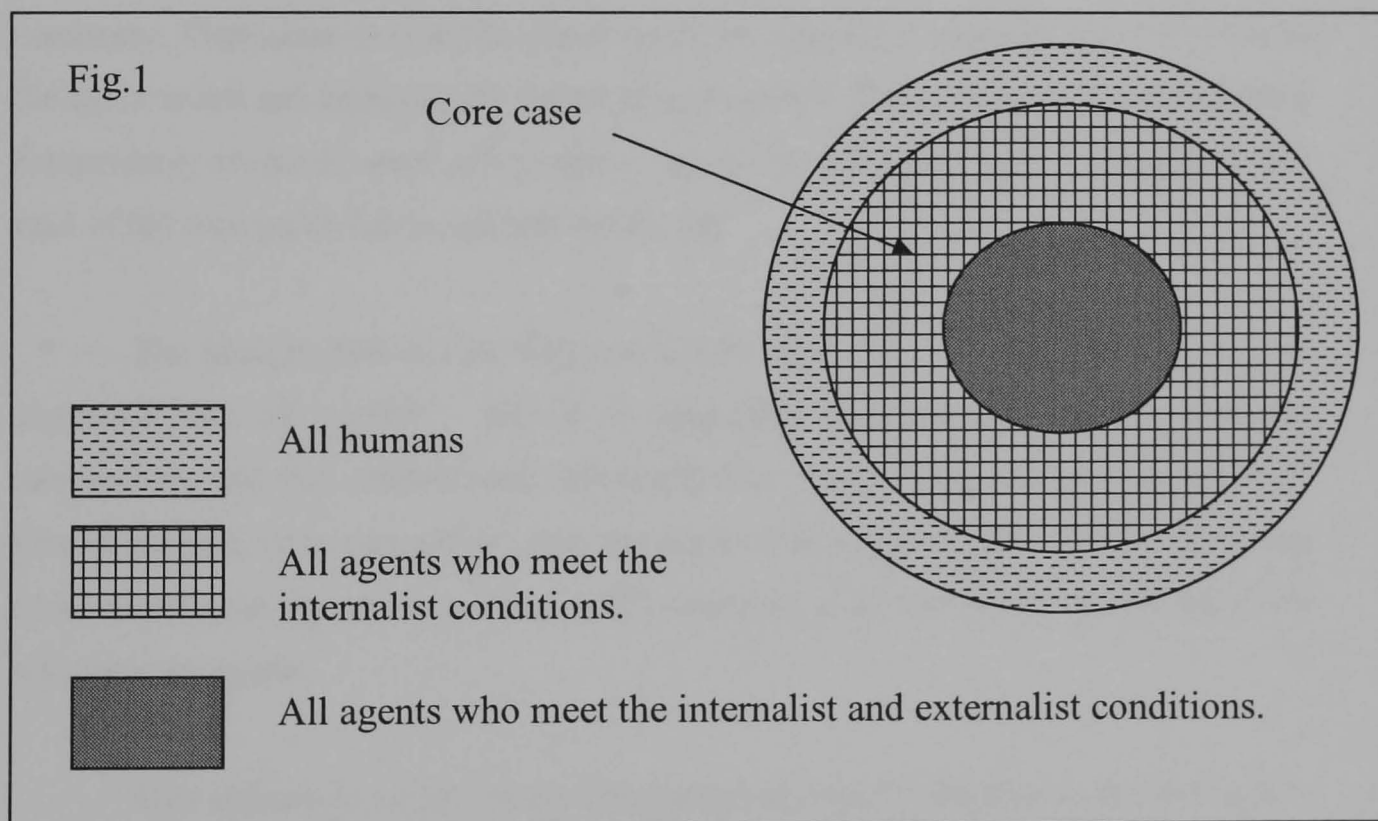
Suppose, then, that we take all the cases that have the structure of the core case, such that a deliberatively competent agent's decision is overridden, for the sake of what another takes to be in her interests. A constitutively relational conception of autonomous agency will only be able to hold that these cases are instances of paternalism in some of the cases that have the structure: namely, those in which the agent stands in the kind of social relations specified.

That is to say, there will be some cases in which the agent meets all of the internalist conditions – manifesting structures, or having undergone procedures, that meet the requirements for deliberative competence, rationality, and so on – but fails to be (according to the relational conditions) autonomous. In such cases, the relational theorist cannot maintain ii) that the intervention is unjustified because it is a violation of autonomy. Because she is not judged to be autonomous, an agent in such a case will not be accorded the normative benefits, one of which is that her decisions and choices should not be overruled, and that the decisions of others should not usurp those of her own.

Thus a constitutively relational conception of autonomous agency does not seem to be appropriately placed to diagnose one of the central claims about what is wrong with interventions in instances that have the structure of the core case – namely, that an agent's autonomy is not respected. This is because such constitutively relational conceptions place externalist social-relational conditions upon autonomous agency –

¹⁷ This is not to say that the social conditions of the agent may not interfere with the agent's meeting of the procedural conditions. Insofar as the social context plays this role, though, we have a different kind of condition, as I spell out in more detail shortly.

requiring that for an agent to be autonomous, not only must the agent be a certain way (according to certain internalist specifications); the world must also be a certain way.¹⁸ This leaves open the possibility of an agent who meets all of the internalist specifications nonetheless failing to be autonomous, and so failing to accrue the normative benefits to which autonomous agents are entitled. In particular, the relational theorist of this kind cannot say, for all instances of the core case that the agent is treated paternalistically and that such interventions are unjustified violations of autonomy. This problem is illustrated in Fig.1, below. In sum, the problem lies in the fact that the core case could fall outside the set of all agents who meet the constitutively relational conditions for autonomous agency. This is the ‘problem of externality’: the relational conditions are externalist. Thus an agent may fail to meet these relational conditions, whilst meeting the internalist conditions relevant to determining whether paternalistic intervention is unjustified.



6.3 Three options

I have shown that, given the structure of an account that incorporates the relational condition (RelAgents), it is possible that the core case might fall into the class of those

¹⁸ On Oshana’s view, this demand is not restricted to the social world; she argues that natural phenomena can also be an impediment to autonomy. She writes: “the number and significance of options made available in the environment of the person are salient concerns for autonomy, and if a person is to be autonomous ... myriad impediments within the environment of the agent must be contained. These include impersonal natural impediments and circumstantial impediments, as well as interpersonal impediments” (2006, *Op Cit.* p.92).

agents who are not, according to the relational account, autonomous. In this section, I set out the options that present as a result. I argue we should take the option that involves rejecting the relational conception of autonomous agency.

Option A: If one wants to maintain that autonomy should protect against paternalistic intervention, such that all such interventions are unjustified because they violate autonomy, one should reject a conception which incorporates ‘externalist’ conditions such as (RelAgents). Thus we reject the claim that the conditions for autonomous agency are constitutively relational. This is option (A), and it is the option that I think we should take. I argue for this by showing that the other options that face the relational theorist are not feasible.

Option B: The relational theorist could nonetheless maintain that the interventions such as those in the core case will be unjustified, even if they are not all violations of autonomy. Thus claim i) could be maintained. This response could maintain that whilst the agent is not autonomous, she nonetheless has some claim – due to her deliberative competence, or due to some other value – not to have her decisions overruled for the sake of her own good. Let us call this option (B).

The thought here is this: why not simply reject the claim – upon which it was suggested that all agreed – that it is autonomy that protects from paternalistic intervention, and that renders such interventions as those in the core case unjustified? One could hold, then, that autonomous agency is relational and maintain that the wrong of paternalism is, say, the overriding of the decisions of deliberatively competent, if not autonomous, agents.

This appears to be one of the strategies endorsed by Oshana, in her defence of her social relational view against objections from liberalism. She notes that the social-relational account she offers “is not comprehensive in the variety of lifestyles it permits [as autonomous].” Thus the concern arises: “there is the risk that the voices on the fringe – especially on the fringe of social and political power – shall be silenced”¹⁹ – these voices may be overridden, and subject, then, to paternalistic interventions.

In response, Oshana suggests that we might maintain that such individuals have de jure autonomy – that is, they have a right to a voice, to make their own choices

¹⁹ *Ibid*, p.99.

without interference. Such a right may well be grounded in the deliberative competences of the agent. But, Oshana claims, “autonomy demands more of a person and of her environment than does de jure autonomy”.²⁰ Thus we can hold that agents who fail to meet the condition set by (RelAgents2) nonetheless have a right to autonomy, in the sense of a right to be left to govern themselves. That the agent is not protected from paternalistic intervention “would follow only if one result of denying [social-relational] autonomy is that de jure autonomy, or the right to be heard, is denied”.²¹ We can then maintain that it is wrong to intervene with the choices of individuals such as agent A, because to do so is a violation of her right to autonomy, whilst accepting that they lack (social-relational) autonomy.

Alternatively, we might maintain, Oshana suggests, that “values such as liberty ... give us grounds to consider limits to state interference”;²² on this basis – an individual’s right to liberty – non-paternalistic policies could be held, whilst still insisting that the individual lacks autonomy. Autonomy in the relevant (social-relational) sense need not be what grounds entitlement to protection from paternalistic intervention – some other value may play this role.

This response has costs, however. The normative frameworks of respect, participation, paternalism and autonomous agency are intimately connected. As such, any one claim cannot be easily detached. Rejecting the claim that it is autonomy that sets the bounds of paternalistic intervention, then, would require rethinking the relationships between autonomy and respect, and those between autonomy and legitimate political participation. Whilst doing so may yield a notion of autonomy that can posit constitutively relational conditions, I submit that it would also rob the notion of autonomy of much normative significance. Not only would autonomy not protect from paternalism; nor would it identify those agents who deserve respect, nor those who are candidates for political participation. No longer would the concept of autonomy play the normatively important roles that motivated feminist theorists to retain and refigure it. ‘Autonomy’ would not ground respect, or protect from paternalism, or entitle agents to a voice in collective decision-making processes. ‘Autonomy’ would designate a class of deliberatively competent agents who stand in certain social relations (some of those

²⁰ *Ibid*, p.100.

²¹ *Ibid*, p.101.

²² *Ibid*, p.118.

agents may not themselves value those social relations). ‘Autonomy’ would lose much of what accounts for its value.

Indeed, Oshana acknowledges that “[relational] autonomy [in her sense] is one among a multitude of values” and that “not everyone will include an autonomous life [namely, one that includes the social roles and conditions she specifies] among the goals that she regards as integral to well-being”.²³ Insofar as autonomy is conceived of as just one value that may or may not be chosen – just one component of well-being that is not necessary for a good life – we see that she departs from a conception of autonomy which plays a role in grounding respect, identifying candidates for political participation and protection from interference: the notion of autonomy that plays these roles extends to all agents, even if they do not take up the entitlement to political participation, or if they adopt conceptions of the good other than those deemed desirable, or other than those that are consonant with the value of substantive independence. A life that is autonomous in Oshana’s sense may indeed have value for those who choose it. But it is not the same core value that a notion of autonomy with its attendant normative benefits has for agents. It may well be that other values also counsel against paternalistic interventions. But we should hold that these are additional, not alternative considerations to those supplied by autonomy. Thus we should not take option B.

Option C: Now, a proponent of the relational conception could argue that the intuitions about the core case are wrong: the relational conception makes the correct diagnosis, and some agents who meet the internalist conditions are nonetheless non-autonomous, and so not protected from paternalistic intervention. For instance, suppose that in the UK, there exists a small subculture of fundamentalist religious adherents. Within this subculture women have a subordinate role, in accordance with the group’s interpretation of the religious doctrines. Women do not have access to education, lack freedom of movement, and do not work outside the house. Violations of the prescriptions face severe punishments. Suppose that an agent endorses the restricted domestic role that she occupies, and does so having carefully considered and researched (despite lack of access to formal education) alternative ways of living. Indeed, she takes her religious commitments to be integral to her identity. Let us suppose also that she meets all the requirements of deliberative competence, then.²⁴

²³ *Ibid*, p.104.

²⁴ In her (2008) Clare Chambers scrutinises the social conditions in which choices for such restrictive lives are made, arguing that liberals have the resources to criticise those practices that,

Nevertheless, whilst she is deliberatively competent in her endorsement of her restricted role, she lacks an adequate range of options, and is not allowed to pursue values other than those enshrined by the teachings of the religion without fear of recrimination, and so on. Thus, according to the relational view, she is non-autonomous. Overriding her choice, then, may be justified. Having argued that her account still has the resources to make sense of the right to non-interference, Oshana then suggests that such a right may not, in fact, be respected in to all those who lack social-relational autonomy:

I believe a case can be made that strong [hard] paternalistic intervention is sometimes needed to preserve autonomy that is threatened by a competent and deserving person's self-regarding conduct, even where the target of the paternalistic gesture has not behaved in ways that clearly permit infringements of autonomy.²⁵

This proposal is justified by the fact that, contra Mill's claim that individuals' interests will best be served by allowing them to make their own choices, it is rather

in a person's interests to be autonomous. ... a person who lacks [relational] autonomy stands on weaker ground when claiming consideration from others and is less likely to be accorded the presumption of consideration human agents are due. The failure of people to decide accurately about their autonomy might offer one reason in favour of paternalistic interferences, even when a person has decided in what he believes is his best interest. The right to [relational] autonomy will be a good that must be preserved, even when it is the aim of the individual to destroy this right in none but himself.²⁶

even if autonomously chosen, are an impediment to equality or justice. Note that I am here concerned only with paternalistic interventions – namely, those justified exclusively by reference to the welfare of the agent. This does not rule out intervention (of some kind) *tout court*. As Chambers argues, some interventions with individual liberty – and usurpations of autonomy – may be justified with reference to the values of equality and justice. See esp. her chapter 6.

That intervention of some kind, based on such other values, is permissible may garner further support from the consideration raised by Danny Scoccia: namely, that our preferences and desires may not be autonomously formed, or at least, are formed in and shaped by particular social contexts. (See Scoccia, D. (1990) 'Paternalism and Respect for Autonomy' *Ethics*, 100 (2), pp.318-334.) If those contexts are unequal, then this may lead to patterns of desire for options that perpetuate that inequality. Thus, whilst an intervention motivated by the agent's welfare alone may be a failure of respect, an intervention on the basis of values such as justice or equality may be called for, and justified. This matter is one that I cannot fully address here.

²⁵ Oshana, M. (2006) *Op. Cit.* p.115

²⁶ *Ibid*, p.116 Note that this claim is strong; to demand that relational autonomy must be preserved suggests that states may be required to prevent individuals from becoming (e.g.) monks or deferential wives. This is implausible, and indeed, is in tension with the claim we have just seen from Oshana – that autonomy is one value amongst others. I believe that this tension is generated by an underlying problem; namely, Oshana's susceptibility to the 'apples and oranges' objection. Whilst attempting to give an account of autonomy in the sense at issue in this thesis

There are two points to be made here, the first pertaining directly to Oshana's argument; the second to its implications. First, then, to Oshana's claim that it is in the interests of an agent to be autonomous, in the social-relational sense that she sets out: it may well be that individuals who stand in the social relations she demands for autonomy have satisfactory lives in which many of their interests are met. But the claim that failing to stand in such social relations erodes grounds for consideration and respect, and destroys an individual's right to autonomy, cannot be substantiated insofar as agents so located maintain a right to autonomy. As we have just seen (and as Oshana herself acknowledges), lacking social-relational autonomy does not entail that one lacks the right to autonomy, in the sense of the right to a voice, and to non-interference in one's choices. Furthermore, we have also just seen that Oshana's social-relational notion of autonomy is hived off from the de jure autonomy that (on the basis of deliberative competence) grounds respect, candidacy for political participation, and so on. Thus there is no reason to suppose that choosing a role that does not involve social-relational autonomy will erode the basis of respect. (Indeed, it is plausible to suppose that intervening with an agent's choice, even for the sake of social-relational autonomy, may be more likely to tend towards the erosion of respect.)

Second, I think that once one acknowledges what the claim that paternalistic intervention in the choices of deliberatively competent agents may be justified entails in practice, it is clearer that we should not accept it. Recall again the agent who endorses her social role of subordination. If this agent is non-(social-relationally) autonomous, and so not protected from paternalistic intervention, then her decisions and endorsements can be ignored and overridden; or she may be coerced into leaving the subculture that she occupies, and certain courses of action – enforcing participation in education programs, preventing her from practicing in or attending religious meetings, say – may be imposed upon her. This seems to me to be unpalatable. To override an agent's decision, especially when that decision concerns matters that she takes to be central to her identity, will likely be experienced as a significant harm. Despite being well-motivated, such an intervention, then, may be significantly detrimental to subjective well-being.

(self-governance such that the agent is entitled to normative benefits), she in fact offers a view of autonomy that is concerned with a particular conception of the good, associated with substantive independence.

Moreover, if the relational theorist is to maintain that agents so positioned may nonetheless retain a right to autonomy, such interventions will also constitute a significant failure of respect. Whilst one might deem the relations that the woman chooses to be detrimental to proper regard for her qua agent, one must also acknowledge that such interventions fail to respect the expressed preferences of the woman. Both the role that she occupies and intervening action appear to threaten the ‘presumption of consideration’ to which persons are due. But that the agent has made sober and careful choices, and that her preferences and subjective sense of well-being will be significantly thwarted, counsels, in my view, against overriding or usurping the agent’s choice.

Paternalistic interventions, fortunately, are not the only ways of responding to such scenarios. Whilst overriding an individual’s choice appears unpalatable, other courses of action might be justified; such as engaging in dialogue, exploring the possibility of egalitarian interpretations of the religious texts, creating support networks for women who may choose to leave the community, and so on.²⁷ Given the availability of other options, to treat individuals as non-autonomous, and to override their decisions in this way seems only to compound any perceived oppression.²⁸ Failing to respect the deliberatively competent agents, and the choices and decisions of these individuals, even when one takes those choices to be gravely mistaken, is not the route to empowering and enabling. Thus we should not take option C.²⁹

6.3.1 A misunderstanding?

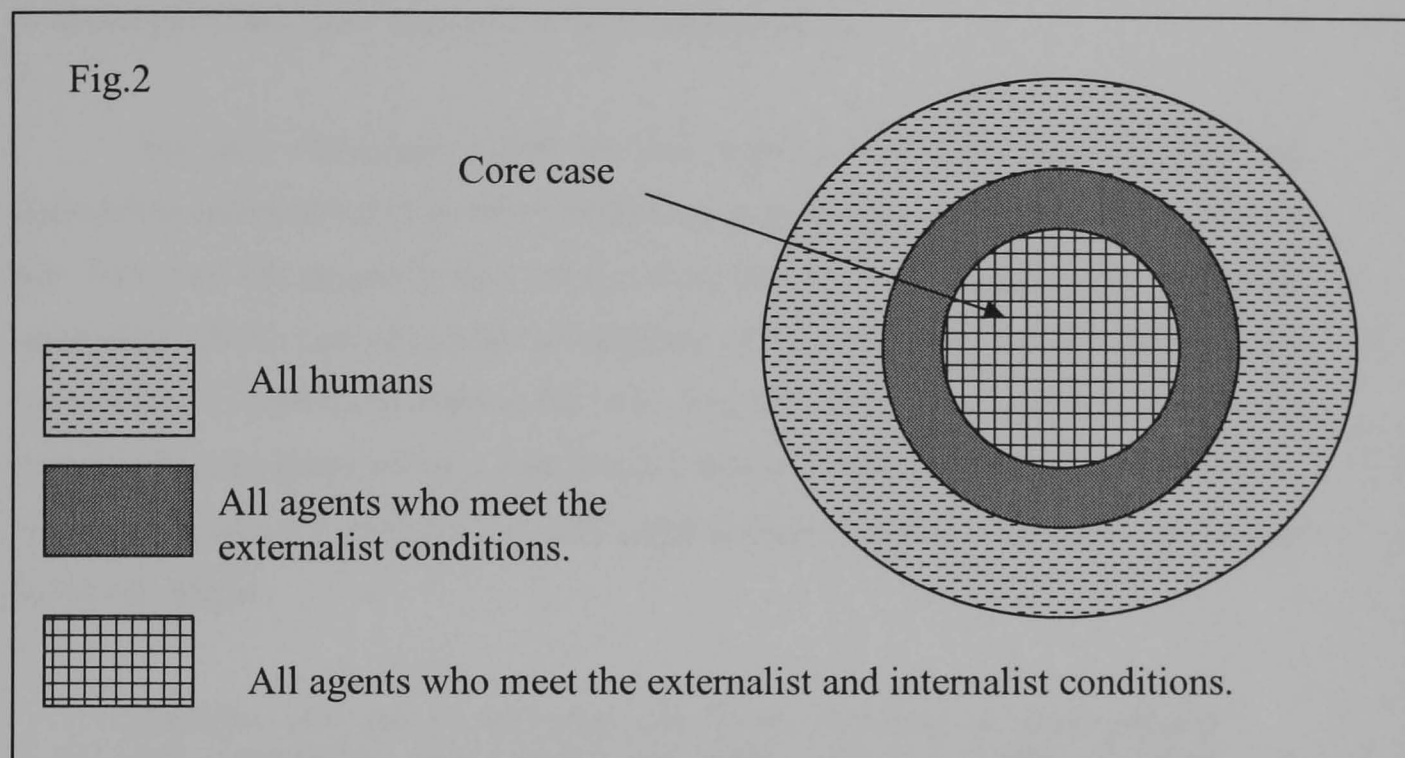
A natural thought would be to claim that the proponent of a relational condition, such as (RelAgents) intends that the relational condition specifies conditions that are necessary

²⁷ Clare Chambers (2008) proposes an ‘equality tribunal’ that might deal with complaints of unequal and harmful treatment in particular cultural or religious groups. She emphasises, however, that interventions would be ‘demand led’: “the tribunal would not intervene in ... proceedings until it had been asked to do so by those concerned” p.136.

²⁸ See Narayan, U. ((1997) *Dislocating Cultures: Identities, Traditions and Third World Feminism*, New York, Routledge) for discussion of how imposition of certain practices (motivated to further the good of a group of agents) can worsen the situation of those whom the imposition intends to help. See also Saul, J. ((2003) *Feminism: Issues and Arguments*, Oxford, Oxford University Press) for discussion.

²⁹ An objection to marshall: why not simply build an anti-paternalist condition into the relational condition – those agents who are autonomous are those who are protected from paternalism? This is an unsatisfactory move that goes beyond the claim that agents who are autonomous are entitled to protection from paternalistic intervention. We should not accept the proposed claim, as a consequence of building a relational anti-paternalistic condition into an account is that agents who are not respected, and are thus paternalistically treated, are not autonomous agents. It should be clear that this claim is problematic, and that rather we should say that such failures of respect are problematic precisely because the agent is autonomous.

for the agent to meet the internalist conditions. Thus it simply cannot be the case that the agent would meet the internalist conditions, but fail to meet the externalist conditions, as set out in figure 2. If this is what the relational conceptions are intending, then there is no reason to think that the relational conception will be ill-equipped to pick out those agents who should be protected from paternalistic interventions.



It should be clear, however, that this is certainly not the manner in which Oshana is intending her condition: she is offering a condition that supplements the internalist necessary conditions, rather than specifies a precondition for their achievement. Moreover, any account which has the structure illustrated in Fig.2 would be plausible only when construed as offering causal conditions for autonomy: certain social conditions might cause the agent to fail to meet the internalist conditions (so claimed Benson). Thus construed, however, the condition cannot also be a necessary condition, insofar as it is plausible that agents can, on occasion, be sufficiently resilient to possess the kinds of deliberative competences necessary for autonomous agency in the face of difficult social conditions. But such resilient characters do exist – as emphasised by Diana Meyers, who points out the high degrees of self-governance exercised by agents in conditions of oppression (as discussed in Chapter 2).³⁰ And as such, these characters pose a counter-example to relational conditions being necessary, constitutively or causally, for the internalist conditions obtaining.

³⁰ Meyers, D. (2000) *Op.Cit.* p.152.

6.4. Conclusions

I have argued that an account that incorporates a constitutively relational condition for autonomous agency is unable to play the desired role in protecting agents from paternalistic intervention. It is worth noting how my objection to such relational conceptions differs from the concern that Christman has.

We saw Christman's concern that accounts of autonomous agency that incorporate constitutively relational conditions cannot play roles such as those I have set out. However, his argument rests on the claim that such conceptions are value-laden, smuggling into the account, in the specification of the social relations that are necessary, some ideal of autonomous agency. We saw, first, the charge that Oshana's conditions in fact require that agents occupy positions that enable substantive independence. Second, we saw Christman's conflation of all constitutively relational conditions with value-laden conditions:

relational accounts of autonomy [are those] according to which persons are autonomous only when they have particular kinds of value commitments.³¹

Christman's concern, then, holds only insofar as it is true that the specification of certain relational conditions imports values into the account. But I have suggested that it is at least conceivable that one could offer a constitutively relational condition that did not smuggle in such ideals. Insofar as this is a possibility, Christman's objection is not telling against all constitutively relational conceptions of autonomy.

My concern, however, is importantly different, and as such poses a problem for all constitutively relational conceptions of the conditions for being an autonomous agent. As I have argued, such conceptions simply incorporate the wrong kind of condition – an 'externalist' condition – to play the role of delineating the problems in cases of paternalistic intervention. It is the capacities of the agents that are relevant to establishing the wrong of paternalism, and as such internalist conditions are required. The set of agents who meet the internalist and constitutively relational conditions (which demands internalist and externalist conditions be met) will be a subset of those who meet the internalist condition. Accordingly some agents who are deliberately

³¹ Christman, J. (2006) *Op.Cit.* at p.139.

competent will not be (relationally) autonomous – the desired claims about unjustified interventions cannot then be generated.

My conclusion, therefore, is that insofar as we want to maintain a conception of autonomy that can play a role in delineating the bounds of paternalism, we ought not accept a constitutively relational conception of autonomous agency. This does not mean, however, that there are not constitutively relational conditions to be discovered elsewhere. Indeed, in the next chapter I will argue for some constitutively relational conditions for autonomous action.

Chapter 7. Relational conditions for autonomous action

In the previous chapters in the thesis, I have argued against the incorporation of constitutively relational conditions for autonomy; more precisely, for autonomous choice and for autonomous agents. Insofar as my arguments hold, are we thereby committed to accepting that the notion of autonomy has merely causally necessary conditions; that social relations might only “affect our free agency accidentally, by virtue of its potential influences on our capacities”?¹ In this chapter, I show that we should not rest content with this claim.

I have separated out examination of the conditions for autonomous choice, for being an autonomous agent, and for acting autonomously. An agent may have certain autonomy-related capacities as an agent; capacities that enable her to engage with the world, to reflect upon her motives, to reason and deliberate. She may exercise these capacities in making her choices and decisions; such decisions being the outcome of self-governed deliberation. But we can now ask whether, when an agent with the autonomy-relevant capacities, and has made an autonomous choice, she is then able to act upon her choices.

In this chapter I show how we can go beyond the claim that there are merely causally relational conditions for autonomy, by setting out ways in which, for the exercise of autonomy – namely, for autonomous action – certain constitutively relational conditions must obtain. I will claim that this is so for the exercise of autonomy in a range of actions. I sketch four kinds of actions for which there are constitutively relational conditions. The first two conditions are drawn from claims in the existing philosophical literature. The latter two conditions I set out are the start of a structure for thinking more clearly about how oppressive social contexts can thwart autonomous actions. I will not provide a complete account of these here, but show the benefits of such an account, once fully specified. I thus provide the foundations and the justification for pursuing further this line of thought.

¹ Benson, P. (1994) *Op.Cit.* see p. 666.

7.1 More on autonomous action

Theorists have expressed a range of intuitions about what autonomy, with respect to action, involves. For example:

- “the autonomous person must be permitted to act on her own behalf”.²
- agents must not only be able to “discern what they really want and care about” – they must also “improvise ways to express their own values and goals”.³
- an autonomous agent must be able to act in such a way as to “reflect, or mirror, the wants desires and cares that [she] reaffirms when attending to them”, and further, her actions must also “accord with them, and especially promote them”.⁴
- being autonomous requires “conceiving of goals and policies on my own and realising them”.⁵

Each of these claims contains reference to an agent’s action, and expresses the thought that an autonomous agent must be able to do something. In particular, she must be effective in the executing of her commitments or intentions. Berlin demands something stronger even than this – that the agent should be able to (at least sometimes) realise her goals. Being autonomous in action, then, is a crucial aspect of autonomy. But despite the intuitions about autonomous action that are expressed, there has been little detailed discussion of the necessary conditions for this aspect of autonomy. Yet it is here, I shall argue, that we can most clearly see that there are relational conditions for autonomy.

In these claims, we find expression of the thought that autonomous agents are able to act, and sometimes to achieve their goals in doing so. Now, there is an enormous variety in human action, and it would be surprising if the conditions for the performance of each action were the same. I next set out two kinds of action for which there are relational conditions; these two views are obvious and plausible, and should need little argumentation. I then set out the two further kinds of action for which constitutively relational conditions are required, and which, with development, provide a useful framework for thinking about the threats to autonomous action in oppressive contexts.

² Oshana, (2005) ‘Autonomy and Free Agency’ in *Personal Autonomy*, Taylor, J.S. (ed.); Cambridge, Cambridge University Press, p.184.

³ Meyers, D.T. (2002) *Op.Cit.* p.19.

⁴ Friedman, M. (2003) *Op.Cit.* p.6.

⁵ Berlin, I. (1969) ‘Two Concepts of Liberty’ in his *Four Essays on Liberty*, Oxford, Oxford University Press, p.131.

7.1.1 Joint actions

Consider all the actions that we intend to do in concert with others: joint actions. Joint actions are clearly relational: we cannot perform them without standing in certain social relations, namely, the relevant social relations with the co-operating individuals. For example, building a house together, or playing in a football team are cases in which agents with shared ends act together. For these joint actions to be performed, certain social relations must obtain. Put simply, another must partake in the joint task of building a house with you, or sufficiently many must play football.

For some joint actions – such as two agents lifting a piano upstairs – it may be true that each agent, A and B, has certain complex beliefs about the intentions of the other (each believes that the other intends to lift the piano), and their beliefs about each others' intentions (A believes that B believes that A intends to lift), and so on (neither would lift the piano if she didn't believe the other would too).⁶

But the social relations that constitute joint action (as Christopher Kutz argues) need not be interactive and communicative to this degree.⁷ For other actions that are joint – such as building the Great Wall of China – the participants may not directly communicate, or have beliefs about the intentions and beliefs of each other. Two builders, at different ends of the wall, may never communicate, and their actions may not be interdependent. It is nonetheless true that the agents are acting together, rather than merely in parallel. On this 'minimalist' understanding of joint action, agents act jointly insofar as they each act with 'overlapping participatory intentions'. That is to say, each must act with the intention to do her part of the collective goal (where the conception of this collective goal is overlapping).⁸ A virtue of this understanding is that

⁶ Tuomela, R. and Miller, K. (1988) 'We Intentions' *Philosophical Studies* 53 pp.367-398.

⁷ The literature on joint or collective action is vast. One of the central issues is how to bridge the gap between what is true of the individual participants, and what is true of the group. This requires specification of the content and form of the intentions of the participants in joint action. See Kutz, C. (2000) 'Acting Together' *Philosophy and Phenomenological Research* 61 (1) pp.1-31, for a good overview of the current debate and an argument for the minimalist position that I mention here. See also Bratman, M. (1992) 'Shared Cooperative Activity' *The Philosophical Review*, 101(2), pp.327-341; Velleman, D. (1997) 'How to Share an Intention' *Philosophy and Phenomenological Research* 57(1) pp.29-50; Gilbert, M. (1990) 'Walking Together: A Paradigmatic Social Phenomenon' *Midwest Studies in Philosophy*, 15, pp.1-14; Gilbert, M. (1992) *On Social Facts*, Princeton University Press, esp. chapter 4; Miller, S. (1992) 'Joint Action' *Philosophical Papers* vol21(3) pp.275-297; Tuomela, R. (2005) 'We-intentions Revisited' *Philosophical Studies* 135(3) pp.327-369, for other views in the literature.

⁸ Kutz, C. (2000) *Op.Cit* pp.4, 10-11.

it can characterise both well coordinated and interdependent joint actions, as well as those in which there is little direct communication and cooperation. Insofar as we accept this unified analysis of joint action, then, we will accept the following constitutively relational necessary condition for such actions:⁹

(RelActionJA) a necessary condition for autonomous joint action is that the agent stands in social relations in which there are other agents with overlapping participatory intentions.

The joint action is in part constituted by social relations; namely, by the relationships between the members of the group each of whom acts with an overlapping participatory intention. For example, suppose an agent intends ‘that she do her part in the collective act of building the wall’. For her brick laying to amount to this, others must also act with suitably similar participatory intentions. Insofar as the agent conceives herself as acting jointly with others (as expressed in her participatory intention), she will not perform the action she intends (a joint one) if there are no other cooperating agents. Whether or not this particular kind of social relation holds, then, can determine whether or not the agent is able to do what she intends. This is more than a merely causal connection between social relations and autonomous action; the social relations are constitutive of the intended action.¹⁰

7.1.2 Social forms and autonomous action

Social relations, broadly speaking, are constitutive of another kind of action. Certain (perhaps very many) actions are only possible within certain social contexts, or ‘social forms’, where by this is meant not only certain attitudes, practices and behaviour, but also the conceptual framework through which certain actions are understood, and the meaning or significance of these actions.¹¹ Joseph Raz writes that:

⁹ If one accepts a different account, there will be different conditions specified; but insofar as the action is joint it will be in part constituted by social relations.

¹⁰ A further way in which actions might be relationally constituted is when one’s projects involve not merely joint actions, but rather are bound up with other agents. This point is made by Susan Mendus, who writes: “a mother will not characteristically see her own aims and projects as distinct and separable from those of her children. Similarly, people who are married will see their projects as theirs collectively”. (See Mendus, S. (1999) ‘Out of the Doll’s House: Reflections on Autonomy and Political Philosophy’ *Philosophical Explorations*, 2(1), pp.59-70, at p.67). That one’s projects are constituted relationally is similar to the claim about joint action – though at a broader level of generality (projects, rather than specific actions).

¹¹ Raz, J. (1986) *The Morality of Freedom*, Oxford University Press. He writes that social forms include ‘shared beliefs, folklore, high culture, collectively shared metaphors and imagination, and so on’ at p.311.

one cannot practise medicine except in a society in which such a practice is recognised. Notice that in principle one may be born into a society with no medical practice or knowledge endowed with an innate knowledge of medicine. One could then cure many diseases, but one could not be a medical doctor. ... A doctor participates in a complex social form, involving general recognition of a medical practice, its social organisation, its status in society, its conventions about which matters are addressed to doctors and which not.¹²

The thought is that the practical identity of ‘medical practitioner’ only makes sense in the context in which a social institution of medicine exists. Extending Raz’s point to action, we can say that medical practitioners are only able to act in that role in a social context in which there is a social form of practicing medicine. There will be some actions – prescribing a course of antibiotics, say – that will only be available in a context in which certain social relations hold; namely, those social relations that enshrine a particular way of practicing medicine. Absent such a social form, the action of prescribing, say, simply cannot be performed. This is not simply a terminological difference, of the same action being differently referred to. Rather, the action itself is different; what it is to prescribe, constitutively, is different from what it is to give drugs outside of the social form of current medical practice.

Raz is clear that this is not only true of practical identities and practices that are formally institutionalised, such as that of medicine. For example:

Bird watching seems to be what any sighted person in the vicinity of birds can do. And so he can, except that would not make him into a bird watcher. He can be that only in a society where this, or at least some other animal tracking activities, are recognised as leisure activities, and which furthermore shares certain attitudes to natural life generally.¹³

Once again, Raz may be understood to be talking of the social constitution of an agent’s practical identity as a bird watcher. Such an identity can only be held in social contexts in which the relevant practices, with the respective roles, are enshrined. But it is again clear that certain actions – particular actions that are part of those practices – will likewise be possible only where the relevant social form is instantiated. For example, actions, such as that of ‘going on a twitching holiday’¹⁴, will only be possible in

¹² *Ibid* p.310.

¹³ *Ibid*, p.311.

¹⁴ ‘Twitch, v.: To participate in the activity of a ‘twitcher’ (sense 4); to watch obsessively for or spot rare birds. Also *trans.*, to observe (a rarity). *slang.*’ At

instances in which certain social relations hold; those in which bird observing is regarded as a hobby. Thus it is clear that there are constitutively relational conditions for the autonomous performance of such actions (although Raz does not explicitly specify, it should be clear that the range will be very broad indeed):

(RelActionSF) a necessary condition for certain autonomous actions is that the agent occupies social relations (or 'social forms') that enshrine the practices in which the action is embedded.¹⁵

Now, this point can be understood as an almost trivial point about the social embeddedness of action: actions are conceived of and performed in social contexts. But the point is of most interest if we consider the possibility of agents moving across contexts in which different social forms are operative; certain actions may be possible in some contexts but not others. Thus the exercise of autonomy will extend to different possibilities in different contexts. For example, a doctor, travelling to an isolated tribal community in which no practice of medicine familiar to us exists, may intend to prescribe a course of antibiotics. But insofar as the social relations do not instantiate the institution or practice of prescribing drugs, writing on a prescription pad will not amount to the action of prescribing. Once again, the social relations play a constitutive, rather than merely causal, role in autonomous action. The presence or absence of certain social forms, in any given context, will enable or thwart an agent's ability to act.

7.1.3 The social constitution of institutional authority

With the notion of social forms in mind, we are now in a position to set out a further way in which social relations might constitute certain actions. Some social forms involve institutional roles, and the agents operating in those roles must be authoritative

http://dictionary.oed.com/cgi/entry/50260726?query_type=word&queryword=twitching&first=1&max_to_show=10&sort_type=alpha&search_id=YDWN-npKf4H-2444&result_place=1
accessed 04/06/08

¹⁵ In earlier chapters, I asked whether there are relational conditions for autonomous choice. Should we accept this condition as necessary for autonomous choice also? I think not: whilst most formulations of choice will specify the objects of choice in the terms of the social forms in which the agent stands, this need not be so. Let us change Raz's example: suppose a doctor from the UK is transported to a community in which the social form of medicine does not exist (an isolated tribe community, say). He might then choose to write a prescription, believing this possible, although the medical practice as it exists in that community does not include giving drugs in this form. We might say that he can choose to do this; but he cannot execute this choice. Strictly speaking, then, this condition is only necessary for autonomous action. To make this claim more precise, work is required on the rational constraints on intention and choice, with regards what an agent believes to be possible (and in what sense). For discussion on such constraints, see Bratman, M. 'Intention, Belief, Practical, Theoretical' (ms.) forthcoming in Timmerman, J. Skorupski, J. and Robertson, S. *Spheres of Reason*, Oxford University Press.

to so act.¹⁶ For example, to arrest, to sentence, to marry, to assess – all these actions can be parts of circumscribed roles, and to perform such actions requires a certain authority. A framework for thinking about this is provided by Alasdair MacIntyre's observation:

What makes [e.g.] an officer an officer is not only that he holds a commission from some duly constituted authority; he must also be recognised as an officer by his subordinates, by his superiors, and by civilians. When this condition ceases to hold, there cease to be officers. Beliefs are partially constitutive of at least some central social institutions and practices.¹⁷

The social constitution of roles and their attendant authority is of import in identifying further constitutively relational conditions for autonomous action. From MacIntyre's comment here, we can identify two necessary conditions for being authoritative in an institutional role.

a. Grounding conditions for authority

First, to be able to act in the role of an officer, an agent must meet what we can refer to as the 'grounding conditions' of authority in a particular role.¹⁸ For example, an officer must 'hold a commission'. Likewise, a doctor must have the relevant qualifications as a doctor, and be employed in such a role. These grounding conditions are intended to indicate aptitude in a particular role. That an agent meets the grounding conditions will often be 'marked' – perhaps by being elected to an official role, such as the commission of officer. This enables others to identify that the grounding conditions are met; this is important, given the other necessary conditions for acting in such roles.

b. Recognition of authority

Second, we can turn to MacIntyre's comment about the constitutive role of belief. The claim is that the beliefs of others – of, say, subordinates, or superiors, or civilians, in the case of the officer – are relevant to the existence of, and engagement in, certain social

¹⁶ To clarify: my aim here is to show how, given the existence of certain institutional roles the engagement in which requires authority, social relations are constitutive of this authority and hence of the ability to act in such roles. This is by no means an endorsement of those roles, or of the authority that attaches to them (some such roles are valuable – teacher, doctor – others are more contentious – military officer, say. I remain neutral on this issue here).

¹⁷ MacIntyre, A. (1973) 'The Essential Contestability of Some Social Concepts' *Ethics* 84 (1), pp.1-9 See p.3.

¹⁸ This terminology is from Hanrahan, R. and Antony L., (2005) 'Because I said so: towards a feminist theory of authority' *Hypatia*, 20(4), pp.59-79. There, they write: "A properly grounded authority is, in effect, an authority who can be expected to make good decisions. She is someone who wields her authority for the proper purposes and in a way that genuinely facilitates those purposes." p. 71.

practices, or social forms. This constitutive role is two-fold; first, MacIntyre appears to be making a general point about the role of officer; the practices that are deemed to fall under this role's jurisdiction.

Second, MacIntyre makes reference to the way that beliefs of others are relevant to the authority of an individual operating in that role. For an individual to count as an officer, he must be recognised as such by the relevant others – subordinates, superiors, civilians, for example. To stand in the role, and to be able to engage in the practices of the role, the officer “must also be recognised as an officer” by relevant others. Meeting the grounding conditions provides the individual with a claim to this recognition; she can hold that others ought to recognise her as authoritative in that role.

19

If MacIntyre is right about this, then social relations will be constitutive of actions that attach to certain institutional roles. When a role requires that agents operating in that role have a certain authority, the actions that attach to that role can be performed only if others recognise the authority to so act. That is, the following relational condition for autonomy can be accepted:

(RelAction1) a necessary condition for authoritative action in institutional roles is that the agent is recognised as authoritative in that role.

c. Treating as authoritative

Whilst MacIntyre's insights here are instructive, I think it is important, however, to go beyond his claims; namely, to note that recognition of an individual (or group) as authoritative is neither necessary nor sufficient for authority, and that rather treatment as authoritative in a role is necessary (and can be with the grounding conditions, jointly sufficient). Beliefs about the practices that make up a role, and about the authority of certain individuals to engage in those practices are an important part of the story; but they are only a part. First, we can say that it is not enough that others recognise the authority of agents in particular institutional roles; such a recognition or set of beliefs will not have any role in shaping social relations unless they are manifested in behaviour. Recognition that an agent meets the grounding conditions for authority is not

¹⁹ I am for now accepting the claim that the belief about the agent's authority is necessary; later we will see that it is not necessary – agents can be authoritative in a role if they are treated as such, and being so treated can be a matter of ‘going though the motions’ in the absence of a belief that the agent in fact is authoritative.

sufficient for the constitution of that authority. It is important that sufficiently many others must treat those agents as authoritative.²⁰ That is to say, sufficiently many others must engage in practices and patterns of behaviour that acknowledge and accommodate the authority of the agent in the institutional role. Accepting this claim, in fact, removes the force of MacIntyre's claims about the beliefs of others constituting their authority. Others may treat an agent as authoritative without in fact believing that the agent has authority. So recognising the authority of an individual is not necessary for the constitution of that individual's authority.²¹ Thus we should revise the constitutively relational condition, from MacIntyre, to:

(RelAction2) a necessary condition for authoritative autonomous action in institutional roles is that the agent is treated as authoritative in that role; that is others must engage in, or be disposed to engage in, patterns of behaviour that acknowledge and accommodate the competence and standing of the agent in that role.

The following example, from Doris Lessing's *The Grass is Singing*, serves to illustrate this framework; the failure of the relational condition (RelAction2) to obtain deprives certain agents of the authority to act in the prescribed institutional roles.

Arrest: Set in the farming communities of South African apartheid, Lessing's novel unfolds the complicated circumstances of a murder. The events leading to the murder, and the relationships between the parties involved, are wholly tainted with the vicious racism of the time. In the following passage Mary Turner's body has been found on the veranda, stabbed. Upon arrival at the scene, the police officers find her husband, Dick Turner wandering madly around the bush, and Moses, the 'houseboy', who offers himself up for arrest. Waiting for the arrival of the Sergeant, the black police officers are unable to take command at the scene:

They snapped the handcuffs on him [Moses] and went back to the house to wait for the police cars to come. There they saw Dick Turner come out of the bush by the house ... He was off his head, talking crazily to himself, wandering in and out of the bush with his hands full of leaves and earth. They let him be, while keeping an eye on him, for he was a white man,

²⁰ Recall that Raz talked, in his discussion of the social form of practicing medicine, of not only the "general recognition of a medical practice", but also of "its conventions". With talk of conventions, Raz goes beyond the matter of mere beliefs about certain roles and the authority to engage in them; we have the notion of convention, or patterns of (mutually acknowledged) behaviour.

²¹ Note that MacIntyre's point about the constitutive role of beliefs with respect to the authority (the practices legitimated) of roles (rather than individuals or groups acting in them) still stands.

though mad, and black men, even when policemen, did not lay hands on white flesh.²²

An attempt to arrest, for the black police officers, would not amount to a legitimate exercise of their authority qua police officer; rather, it would be a violation, an assault. The framework set out enables us to make sense of this: whilst the officers meet the grounding conditions for authority in the role of police officer, they are not, under South African apartheid, treated as having the authority to so act. The role of police officer, in this context, is itself racially stratified; the officers are not treated as authoritative when it comes to exercising police powers over white citizens. That the role of police officer is socially ‘carved up’ in this way should not be surprising, given that the rest of apartheid South Africa was also structured along hierarchical race lines. Black police officers were treated as authoritative in only some of the practices that white police officers were. The failure to be treated as authoritative undermines the authority of agents to operate in certain institutional roles. If we accept MacIntyre’s claims about the social constitution of authority, then for actions that require authority of this kind constitutively relational conditions for the exercise of autonomy obtain: one cannot so act unless one stands in social relations in which being treated as authoritative in that role is assured. This kind of relational condition will hold for all of those actions that are undertaken under the remit of an institutional role (actions of military officers, police officers, doctors, teachers, social workers, factory workers, office assistants, stock brokers, and so on and so forth).

7.1.4 The social constitution of social authority

I now want to sketch out how the kind of relational condition identified in the previous section might provide us with a way of thinking about more everyday, non-role specific actions. I’ll suggest that, if fully developed, this framework can yield a fruitful way of thinking about how oppressive social relations might thwart the exercise of autonomy in subtle ways.

Suppose that operating as an autonomous agent is understood to be structurally akin to operating in an institutional role. We might say that certain actions fall under the remit of the ‘role’ of autonomous agent; performing those actions requires both that an agent has some kind of authority, meets some grounding conditions for that authority,

²² Lessing, D. (1950) *The Grass is Singing* London, Harper Perennial, p.12.

and is treated by others as authoritative in that role. If all this can be plausibly filled out, then there will be a constitutively relational condition for exercising autonomy in those actions; agents must be treated as authoritative qua agent.

In order that this be more than mere speculation, it is required that something be said about the following three aspects: the grounding conditions for authority; the kind of authority at issue and the appropriate treatment thereby; and the actions for which this authority is required. I'll very briefly say something about how I envisage each of these matters might be more fully substantiated. Finally, I will give some examples which, I hope, add support to this strategy by making plausible the claim that a kind of authority is indeed needed for actions that are not 'role-specific' or institutional.

This is what Louise Antony and Rebecca Hanrahan have to say about the role of authority in interpersonal interaction:

authority is a thoroughly social phenomenon, and earns its place in a properly ordered society by making possible various forms of cooperation, forms that are necessary if we human beings are to exploit our unique combination of capacities for knowledge and communication.²³

The idea here is that communicative and cooperative action requires relations of authority. We've seen that authority in institutional roles is socially constituted; much cooperation and communication will be served by these formal relationships of authority. But much communication and cooperation occurs outside of any institutional role; thus perhaps this kind of interpersonal interaction may be understood as part of what is distinctive about the 'role' of agents, for which a kind of authority is required.²⁴

If authority qua agent is required for operating in this role, how might we understand this authority? A look to an intrapersonal analogue may be instructive. Benson understands the notion of intrapersonal authority as integral to the conditions for being an autonomous agent. The kind of self-regard involved in the process of self-authorisation that Benson emphasises is one according to which the agent has "ownership with regard to [her] conduct", and this amounts to the agent having a sense

²³ Hanrahan, R. and Antony L, (2005) *Op.Cit.* see p. 69.

²⁴ Indeed, Miranda Fricker gives a detailed account of the social relations that shape attributions of epistemic authority that characterise our practices of knowledge transmission. See Fricker, M. (2007) *Op.Cit.*

of her “social competence or worth ... regard for [her] abilities and social position”.²⁵ Autonomous agents must “secure in their own minds their regard for their competence and worthiness to speak for themselves”.²⁶ First, there is the agent’s sense of her competence; second, there is the agent’s sense of her worth. This is an evaluative matter, pertaining to the agent’s status qua agent, among others.

A plausible interpersonal analogue, then, holds that being regarded as authoritative qua agent is a matter of being recognised firstly, as having certain competences as an agent; those competences in planning, reasoning, deliberative reflection and so on, as possessed by autonomous agents.²⁷ But, second, there is an evaluative component, which involves being regarded as worthy, in Benson’s terms.

We might say that this involves a descriptive judgment that captures the grounding conditions for the authority; they ground the agent’s claim to be regarded as worthy qua agent. If agent A makes such a judgment about agent B, she should, insofar as she is rational, take up a certain regard for that agent: as worthy of respect. Importantly, an agent, A, who makes this judgment about agent B should treat agent B as authoritative as an agent; they should act so as to accommodate and acknowledge the competence and worth of the agent, in particular with respect to her competence and worth to engage in communicative and cooperative actions.

I’ve added some detail to the sketch of this way of understanding autonomous action as relationally constituted. In summary, (some of) the distinctive actions of the ‘role’ of an agent are communicative and cooperative; for these actions, an agent must be treated as authoritative qua agent; being so treated involves being treated as competent and worthy to perform these communicative and cooperative actions; this authority has as its grounding conditions the competences and capacities that constitute autonomous agency. If this view is right, then a constitutively relational condition for autonomous action will be:

²⁵ Benson, P. (2005a) *Op.Cit* pp.110-111.

²⁶ *Ibid*, p.115.

²⁷ See Bratman, M. ((2005) ‘Planning Agency, Autonomous agency’, in *Personal Autonomy* Taylor J.S. (ed.) pp33-57) for the claim that planning is integral to our agency. The claim that our capacities for reflection and reasoning are integral to our agency is well-established. See Kant *Op.Cit.* for a statement of this claim; see Frankfurt, H. (1971) *Op.Cit.* for a more recent statement of the thought.

(RelAction3) a necessary condition for certain communicative and cooperative autonomous action is that the agent is treated as authoritative qua agent; that is, others must engage in, or be disposed to engage in, patterns of behaviour that acknowledge and accommodate the agent's competence and worth as an agent.

Is this plausible? I offer the following examples by way of support for this claim. The following cases, I contend, are ones in which agents are unable to exercise their autonomy due to the failure of the relational condition (RelAction3) from obtaining.

Registration: In the following, we see an agent who is unable to register to vote. This, I submit, is because he is not treated as authoritative as an agent. The action of registering to vote is not an action that is 'role-specific' in the way that arresting or prescribing is; nonetheless a kind of authority is required for its performance. Under Jim Crow segregation in the US, this kind of authority was denied a whole class of citizens: many black citizens – as well as poor uneducated white citizens – were unable to register to vote – even after the legal disenfranchisement of black citizens officially ended. In *Remembering Jim Crow*, Leon Alexander describes his experiences, in 1910, of attempting to register to vote:

I walked in and the registrar was in there. He started washing his hands. He washed his hands until two white people came in and he dried his hands off and came and waited on them. Then he went right back to washing his hands again. Finally he came over and just like this he said: "What you want boy?" I said, "I wants to register to vote." So he got out a registration form and laid it out there before me ... He came back and looked at it, balled it up and threw it in the wastebasket. All he said was, "You disqualified, you didn't answer the question".

...Somebody asked me where I had been. I told him I had been over the courthouse trying to register to vote. So one of them asked me: "Well did you have any luck?" I said, "No, they didn't even read the thing. He just picked it up and balled it up and throwed it in the waste-basket and told me I didn't pass.

... So that's how my first voting experience went because this guy had no intention of registering me, not only no intention of registering me, he had no intention of registering any black to vote.²⁸

This action of registering to vote, then, seems to be an action that one must be treated as an authority to perform. It is a cooperative action that is only available to those treated as competent and worthy as agents. And it is certainly the case that Leon Alexander is

²⁸ *Remembering Jim Crow*, Chafe, W.H., Gavins, R., Korstad, R. (eds.) 2001, New York, The New Press, pp.277, 280. Detailed discussion of disenfranchisement and segregation laws can be found in Klarman, M.J. (2004) *From Jim Crow to Civil Rights: the Supreme Court and the Struggle for Equality*, Oxford, Oxford University Press.

not treated as having the authority to so act, and thereby cannot. This case seems to be well explained by the framework I have spelt out, and provides illustrative support for it.

Refusal: There has been much recent attention to women's ability, or lack thereof, to refuse in sexual contexts.²⁹ Rae Langton has claimed that in such a context, norms of sexual communication may be operative such that "'no' means 'yes'": refusal is understood as coy permission, say (her concern is with pornography's role in bringing this about).³⁰ In such contexts, women's attempted refusals may not amount to refusals.³¹ However, there are other contexts in which women's refusals appear to be 'unspeakable' – and these in such contexts it is clearer that there are no norms that might muddy the matter of what is intended in the utterance of 'no'.

Women's attempted refusals of medical treatments in pregnancy and birth is increasingly documented. On a website that enables women to share their experiences of traumatic treatment they received in giving birth, there are a number of distressing reports of utterances of 'No' which fail to count as refusals:

I was repeatedly examined vaginally, even when I told them 'NO', [but] even when I said, "it hurts" even when I said "it burns" and they told me "No you can't feel that, it doesn't burn, you don't have nerve endings that far up". The doctor used an amniohook to break my waters without my knowledge or consent. ... Then when my son was born, they clamped and cut his cord, administered pitocin (again without my consent, I had already stated "No pitocin" during labour, I did not consent to the use of pitocin at all). And he began pulling at my placenta within minutes of my son's birth,

²⁹ My thinking about such cases in terms of autonomy was prompted by Westlund, A. 'Women's autonomy in conditions of oppression' (unpubl. ms.).

³⁰ See Langton, R. (1993) 'Speech Acts and Unspeakable Acts' *Philosophy and Public Affairs* 22 (4) pp.293-330; Langton, R. and Hornsby, J. (1998) 'Free Speech and Illocution' *Journal of Legal Theory*, 4, pp.21-37; Langton, R. and West, C. (1999), 'Scorekeeping in a Pornographic Language Game' *Australasian Journal of Philosophy*, 77 (3), pp.303-319. For discussion of some of the problems facing these views, see Maitra, I. (2004) 'Silence and Responsibility', *Philosophical Perspectives* 18 (1), pp.189-208; Jacobson, D. (1995) 'Freedom of Speech Acts? A response to Langton' *Philosophy and Public Affairs*, 23(1) pp.64-79; Wieland, N. (2007) 'Linguistic Authority and Convention in a Speech Act Analysis of Pornography' *Australasian Journal of Philosophy*, 85 (2), pp.435-456.

³¹ Harris writes of the legal contexts in which women were denied the authority to refuse: in the US, even after the civil war, black women were 'unrapable' in law. See Harris, A. (1993) 'Race and Essentialism in Feminist Legal theory' in Weisberg, D. K. (ed.) *Feminist Legal Theory: Foundations*, Philadelphia, Temple University Press, pp.349-358. Her observation is offered as a criticism of Catherine MacKinnon's analysis of 'woman' as 'rapable'. See MacKinnon, C. (1987) *Feminism Unmodified*, Harvard University Press.

I screamed at him “what are you doing, stop, I could haemorrhage”, he said “no I gave you pitocin.”³²

This case provides some support for the claim that authority is required to perform communicative actions such as refusal. One way of understanding this case is as one in which the action of refusing medical treatment requires authority – and, importantly, being treated as authoritative. The woman, who identifies herself as ‘EB’ utters ‘no’. She attempts to refuse both examinations and the administering of drugs. (It is absolutely clear that the issue here is not one of misunderstanding; there is no suggestion that the context is one in which, as is suggested in the case of sexual refusals, ‘no’ means ‘yes’.) But nonetheless, EB’s refusals are not effective. Now, it might be that her utterances are recognised as refusals, but overridden, and hence she refuses but fails to achieve the effects she intends. However, in medical contexts, the doctrine of informed consent is presumed to hold (namely, that the patient’s informed consent should be secured before medical interventions are performed).³³ Given this, if the refusal is recognised as a refusal, we must understand the medical staff to be boldly and publicly flouting this doctrine. This may counsel against this understanding of the situation. An alternative way to understand what is going on here is in line with the framework I set out above. We can say that EB is not treated as an authority – as a competent and worthy agent – so that her utterance does not thereby count as a refusal. This is a plausible understanding, given Susan Bordo’s observations on the regard in which pregnant women are held:

in this culture the pregnant woman ... comes as close as a human being can get to being regarded, medically and legally, as ‘mere body’, her wishes, desires, dreams, religious scruples of little consequence and easily ignored in (the doctor’s or judge’s estimation of) the interests of foetal well-being.³⁴

I hope these examples have provided some support for the claim that a kind of authority is required for the performance of those communicative and cooperative actions that

³² Posted on TrueBirth.com, Swimm, D. ‘More than a Traumatic Birth’, comment from EB on February 11th, 2008 5:31 pm. <http://www.truebirth.com/2008/02/10/more-than-a-traumatic-birth/>. The experiences of violation and trauma in pregnancy and birth are sometimes referred to as ‘birth rape’. See also <http://www.birthtraumaassociation.org.uk/index.htm>, and Reed, A. ‘Not a Happy Birthday’ *The F word: Contemporary UK Feminism*. http://www.thefword.org.uk/features/2008/03/not_a_happy_bir.

³³ The status of this doctrine is itself contentious. For a detailed exploration, see Manson, N.C. and O’Neill, O. (2007) *Rethinking Informed Consent in Bioethics*, Cambridge, Cambridge University Press.

³⁴ Bordo, S. (2003) ‘Are Mothers Persons?’ in her *Unbearable Weight: Feminism, Western Culture and the Body*, California, University of Berkeley Press, pp.71-97 See p. 76.

have central role in much of what we do as agents. More detail is required in filling out the details of this framework. I believe that it is worth pursuing this line of thought, and motivate this further by setting out, in the following, the virtues that could accrue to this relational understanding of autonomous action.

7.2 Virtues of a relational conception of autonomous action

I'll start first by considering the specific virtues that attach to the particular proposal that I have outlined. I will then set out the more general benefits of giving constitutively relational conditions for autonomous action; namely, that the problems I set out earlier in the thesis can be avoided. Finally, I will briefly suggest some further points of interest to be pursued if my proposal is fully developed.

7.2.1. Diagnosis and remedy of thwarted autonomy

The examples I have given present cases in which an agent's exercise of autonomy is thwarted by oppressive social contexts. I have suggested that attending to the socially constituted authority that is required for performing certain actions might be fruitful for making sense of such cases. If fully developed, we will have a framework for making sense of how oppressive social relations can undermine an agent's authority to act. In order to further develop the account, it will be important to attend to the kinds of social (and sometimes legal) norms that can undermine authority to engage in cooperative and communicative action. Recall that we said that sufficiently many others must treat the agent as authoritative: it will also be important to attend to the matter of how many is 'sufficiently many', when it comes to treating an agent as authoritative – on occasion or more generally.

I think this strategy will be useful not only for diagnostic purposes – for identifying occasions on which an agent's autonomy in action is thwarted – but for remedial tasks also. For example, we can think about how an agent's authority to act might be secured. One way is presented by Leon Alexander's continued account of his voting experience. He recounts how his initial failure to register was not the end of the matter, as he returned to the registrar with some friends:

...I walked in, and they were standing in the hall when I walked up there. Same guy. He saw me and his hands got dirty again, so he went to the sink to wash his hands. So finally Tom Crawford and Ed Bean walked in,

and he dried his hands off and went to wait on them, and they told him, said, “He was here before we were. Why don’t you wait on him?” He said, “Well, hell, I done disqualified him”. Then Tom Crawford said “Let me see that paper that you threw in the wastebasket there”. The registrar said, “who is you?” He still ain’t reaching at the paper, so Crawford told him, “I’m the district representative for the United Mine Workers, and I want to see that paper. I want to see what grounds you put on there that you disqualified him on”

...He still wouldn’t give him the paper ... and Ed Bean went around the counter and got the paper out of the wastebasket and he smoothed it back out. ...Bean said “I’m going to give it to [the Governor] to show him what you are doing to blacks here in Jefferson County”. ... So the registrar got another form and handed it to him. I filled it out. There wasn’t any difficult questions on it.³⁵

His (white) friends, Bean and Crawford, are able to take up Alexander’s cause and force the registrar to treat him as entitled to register. The manner in which the registrar engages with Alexander is imperfect; it would be naïve to suppose that the registrar’s attitude shifts to one in which he regards Alexander with respect. But his going through the motions of the treatment to which authoritative agents are entitled is enough to enable Alexander to act as he intends, which is a start. Thus we see how agents perhaps individually, though more likely collectively – may be able to assert their authority, demanding (for each other) the treatment to which they are entitled. Insofar as we accept the claim that authority is required for such cooperative and communicative actions, it will be important to attend to the ways in which agents can assert their authority and entitlement to be treated as an autonomous agent.

7.2.2 The positive role of respect, and a critique of substantive independence

In chapter 1, I set out the normative benefits of autonomy, one of which was the entitlement to respect. I alluded to Darwall’s notion of recognition respect. This kind of respect means giving appropriate weight to the fact that the agent is a person – taking the fact they are a person to constrain one’s action in certain ways. The claim is that “various ways of regarding or behaving towards others, and social arrangements that encourage those ways, are inconsistent with the respect to which all persons are entitled.”³⁶ If we are able to develop an account of the socially constituted authority required for a range of interpersonal actions, we will have a view according to which being treated as a competent and worthy agent has a role in enabling action. A view such as this would go beyond Darwall’s claims about the role of recognition and

³⁵ Alexander, L. *Op.Cit.* p.279.

³⁶ Darwall, S. (1977) ‘Two Kinds of Respect’ *Ethics* 88 (1) pp36-49 See p.36.

respect. Darwall emphasises that taking up this respect and holding persons in this regard is something we ought to do when we are deliberating about what to do – how, when our actions are to impact upon another, we ought to take that fact into account. On his view, treating another as having status or authority qua agent should constrain our actions; it has a negative role.

However, if my proposal is followed through, we may see that having a certain respect for another – treating them as authoritative qua agents – is relevant not only to constraining our engagement with others. Rather, according to the proposal, it is important in enabling the actions of others. If treating as authoritative in part constitutes that authority, then treating another as such will require not only giving weight to factors that constrain one’s own actions; it will also require taking up a positive role of acknowledgement and accommodation of the communicative and cooperative practices of agents. This has significance if we remember Friedman’s concern (from Chapter 1) that whilst we might want to criticise those who afford too much value to substantive independence, the notion of autonomy could not generate the basis for such criticism. She holds that:

a critique of substantively independent behaviour will have to be based on something other than the ideal of autonomy. We cannot fault autonomy theories for failing to do what might lie beyond their proper scope.³⁷

Already, the relational conditions for autonomous action show that there is reason to be sceptical of the value of substantive independence – if such independence means failing to stand in the social relations that enable joint action, say. Strictly speaking, an agent is dependent upon another if the performance of their action requires the cooperation of another – of an agent with the cooperative participatory intention, say; or perhaps, with an agent who treats one as autonomous. If we can develop the relational condition I have proposed, then we will have further autonomy-based reason to doubt the value of substantive independence. For substantive independence will preclude standing in the kinds of relations of dependency required for acting jointly.

³⁷ Friedman, M. (2003) p.93.

7.2.3 Benefits of locating relational conditions with autonomous action

I now turn to consider the more general virtues of locating the constitutively relational conditions for autonomy at the point of its exercise – namely, autonomous action.

a. The problem of false negatives

In chapter 4, I argued that those accounts which incorporated constitutively relational content into an account of autonomous choice faced the problem of false negatives; namely, choices that are intuitively autonomous came out as non-autonomous on the relational view. I want to here show that, first, incorporating relational content into an account of autonomous action does not face the problem of false negatives. Second, that with respect to my specific proposal, there may in fact be virtues to this approach.

First, in identifying relational conditions for certain kinds of action, we are able to identify problem cases in which the absence of certain social relations prevents the agent from acting as intended: the joint action of (e.g.) playing football cannot be performed; or the doctor cannot prescribe, properly speaking; or the black police officers lack the (counterfactual) power to arrest white citizens. Given that the relational condition is posited to make sense of these failures, the problem of false negatives has no application here (there is not an action that is wrongly diagnosed as non-autonomous; the agent's intended action is not performed at all).

There is, however, one kind of case in which the claim that an action has not been performed has been met with objection. The objection may be understood as a particular instance of the problem of false negatives. This kind of concern has arisen with regards the case of failed refusals in sexual contexts, as mentioned above. There, I suggested that we might make sense of refusals, and the failure to perform a refusal, in terms of the authority required to refuse. Perhaps, I suggested, there are contexts in which agents are not treated as having the authority to refuse, and so cannot.

Previous discussion of such cases has faced the following worry, as expressed by Alexander Bird: in a sexual context in which a woman utters 'no' we should want to say that, insofar as the 'no' is not heeded, a rapist "has committed a cognitive error of a morally heavily loaded kind. But all this can only be said if there is a refusal".³⁸ The concern is this: unless we can say there was a refusal, we cannot generate the desired

³⁸ Bird, A. (2002) *Op.Cit.* p.7. His suggestion is that we should instead say that there is a refusal, but it is understood to be insincere.

claims about the degree of culpability that falls with the man who fails to heed the utterance of no.

If we develop the outlines of the authority based account, as I have suggested, I believe this kind of problem can be adequately addressed. Insofar as an agent cannot perform a refusal because she is not treated as autonomous – as competent and worthy to perform that communicative action – we can identify an “error of a morally heavily loaded kind”. To fail to treat someone as having the authority to which they are entitled is a significant moral failing. There is no reason, on this count at least, to insist that a refusal was performed.³⁹

b. The problem of value

In Chapter 4, we also saw John Christman’s concern that constitutively (rather than causally) relational conditions – conditions according to which “being autonomous means standing in proper social relations to surrounding others and within social practices” – imported values.⁴⁰ But to make such a demand is essentially to demand that the agent hold certain value commitments; requiring that the agent stand in certain social relations amounts to the demand that she not choose, or exercise her agency, in accordance with certain values.⁴¹

Relational theorists who adopt substantive or relational conditions, according to Christman “are in fact supporting a conception of autonomy which is an ideal of individualised self government”.⁴² Indeed, he characterises relational conceptions of autonomy as those “according to which persons are autonomous only when they have particular kinds of value commitments”.⁴³ The problem with value-laden conceptions of autonomy, he claims, is that

to say that she is not autonomous implies that she does not enjoy the status marker of an independent citizen whose perspective and value

³⁹ This is not to say that there are no complications for apportioning blame for (e.g.) sexist acts in conditions of gender oppression. For discussion, see Benson, P. (2004) ‘Blame, Oppression and Diminished Moral Competence’, in *Moral Psychology: Feminist Ethics and Social Theory*, DesAutels, P. & Walker, M.U. (eds.) Rowman & Littlefield, pp.183-200

⁴⁰ See Christman, J. (2004) *Op.Cit.*, at p. 158.

⁴¹ This connection was made clearer with the discussion of Oshana’s view, in chapter 6.

⁴² *Ibid*, p.151

⁴³ Christman, J. (2006), *Op.Cit.* p.139

orientation get a hearing in the democratic processes that constitute legitimate social policy.⁴⁴

The thought is that because autonomy comes with these normative benefits – regard as an agent whose views and choices are worth consideration, and who can participate in political processes (and, as we shall once again see, be protected from paternalism) – it is inappropriate to determine who is autonomous on the basis of the values to which each subscribes. A result of doing so would be that:

authority is circumscribed to exclude voices who are otherwise competent and authentic in ways that procedural accounts of autonomy require.⁴⁵

A relational conception of autonomy that is value-laden, then, classifies agents as non-autonomous for the wrong kind of reason.

Now consider the kind of conditions that I have (in some instances tentatively) identified as constitutive of certain actions:

(RelActionJA) a necessary condition for autonomous joint action is that the agent stands in social relations in which there are other agents with overlapping participatory intentions.

(RelActionSF) a necessary condition for certain autonomous actions is that the agent occupies social relations (or ‘social forms’) that enshrine the practices in which the action is embedded.

(RelAction2) a necessary condition for authoritative action in institutional roles is that the agent is treated as authoritative in that role; that is others must engage in, or be disposed to engage in, patterns of behaviour that acknowledge and accommodate the competence and standing of the agent in that role.

(RelAction3) a necessary condition for communicative and cooperative action is that the agent is treated as authoritative qua agent; that is, others must engage in, or be disposed to engage in, patterns of behaviour that acknowledge and accommodate the agent’s competence and worth as an agent.

These conditions are clearly constitutively relational. But they do not incorporate any values into the account of autonomy; they merely describe the kind of social conditions that must obtain if agents are to perform certain actions (it remains an open question

⁴⁴ Christman, J. (2004), *Op.Cit.* p.157

⁴⁵ *Ibid*, p.157

whether the actions performed are valuable or not). Thus we can see that Christman's conflation of relational conditions with value-laden conditions is mistaken.

Suppose we develop the authority condition I proposed, in outline, for communicative and cooperative actions (RelAction3). Might one object that such a demand imports some ideal or other? One might hold that this condition imports value by demanding that agents hold, and act according to, certain values; namely, it demands that agents value the competence and authority of other autonomous agents. If this is so, and if (as I acknowledged) value laden conceptions are problematic, then this may be a reason not to further pursue this line of inquiry.

Such a condition does not 'smuggle in' such values, however. Whilst it may be true that in order to perform certain actions, an agent must be treated as competent and authoritative, the relational account itself does not insist that agents hold this value in order to be classed as autonomous. (By way of contrast, recall that the substantive views we considered (in Chapters 2-4) demanded that agents valued substantive independence.) It rather describes a certain relation between action and social conditions: if certain actions are to be performed, then certain necessary conditions, involving the regard and treatment of others, must obtain. To perform certain cooperative autonomous actions, the agent must be treated as competent and worthy as an agent. This is not yet to require of all agents that they do treat others as competent and worthy (though I do believe that they ought to do so) in order to be, or choose, or act autonomously. It is rather to identify a relational condition that can enable or prevent other agents from acting autonomously. The ideal of mutual regard and treatment accordingly, is a moral demand, which holds independently of any considerations pertaining to the enabling role of such regard.⁴⁶ The relational condition I have tentatively suggested, then, is purely descriptive, and does not require that agents subscribe to certain ideals or values (beyond those generated by the basic norms of morality). Thus concerns about smuggling in values pose no barriers to the further development of this view.

c. The problem of externality

Christman claimed that the value-laden nature of relational conditions meant that a conception of autonomy that incorporated them could not play certain key roles.

⁴⁶ Of course, this claim requires argumentation, but I cannot give such an argument here. Korsgaard, C (1996a, 1996b) *Op.Cit.* attempts to argue for this claim.

Namely, such a conception could not stand in the normative frameworks of protecting agents from paternalistic intervention. Recall that an agent who meets all the procedural requirements may choose to occupy social relations that do not accord her, for example, a certain range of options (this was one of the demands from Oshana). Such an agent is not autonomous, according to a constitutively relational account of autonomous agency. This means that she is not accorded the normative benefits that attach to autonomy. In particular, it means that

despite her authentic, competent, and ‘sober’ acceptance of such situation (by hypothesis), her lack of relational autonomy – should we accept that view – would allow other agents and representatives of coercive social situations to intervene to relieve her of this burden and to restore her autonomy (at least in principle).⁴⁷

Now, this concern, for Christman, was parasitic upon the claim that relational conditions incorporated value-laden content. I have suggested, however, that we can identify constitutively relational conditions for autonomous agency that are not value-laden. However, that the conditions are value-neutral does not mean that this objection is immediately avoided.

In Chapter 6, I argued that independently of whether any values were imported into a conception of autonomous agency, constitutively relational conditions are simply the wrong kind of condition to do the work in identifying those agents who should be protected from intervention. Whilst internalist conditions explained why agents should not be subject to coercive intervention, relational conditions incorporate externalist conditions. Thus we saw that such a relational conception could only pick out a subset of those agents who, intuitively, should be protected from paternalism. This was the problem of externality. This objection is targeted against relational conditions for autonomous agency. But might it undermine the project of developing further relational conditions for autonomous action also?

It should be clear that it does not. This is because insofar as we accept (some of) the conditions set out above, we are accepting conditions that are constitutive of autonomous action. That is, when such relational conditions are required and fail to obtain, we do not have an action that fails to be autonomous (and hence not protected

⁴⁷ Christman, J. (2004) *Op.Cit.* p.157.

from intervention, and so on). Rather, the action has not been performed at all.⁴⁸ It simply makes no sense to claim that, when such a condition is not met, the agent's action is non-autonomous, so can be interfered with. There is no problem combining this constitutively relational conception with an internalist account of the conditions for being an autonomous agent. Such an account could happily play the role of identifying those agents who are entitled to protection from paternalism. Moreover, I anticipate that further development of the relational condition I suggested would serve only to underline the importance of interpersonal respect and regard for competent and authoritative agency.

7.3 Summary

In the previous parts of the thesis, I set out what constitutively relational conditions might look like, and I argued against incorporating such conditions into a conception of autonomy. In this chapter, however, I outlined out some of the ways in which social relations can constitute autonomous action, and thus are necessary for the exercise of autonomy. We've seen, then, that social relations are connected to autonomy by virtue of more than their potential to causally impact upon the agent's capacities. Rather, the presence or absence of certain social relations can enable or prevent an agent from exercising her autonomy.

In addition to setting out some obvious ways in which actions are socially constituted, I also suggested a framework for making sense of how oppressive social contexts may thwart autonomy. I suggested that we might see autonomous agency as involving a kind of authority to engage with others, in a way structurally similar, though substantively different, to the authority that attaches to various institutional roles. Much work is to be done in filling out the details of this kind of relational condition. However, I hope to have provided good motivation to further pursue this line of thought, by showing the virtues that I envisage attaching to this view, once fully developed.⁴⁹

⁴⁸ Naturally this is a point at which matters of action individuation arise. The important point for present purposes is that the intended action cannot be performed. I cannot here resolve the tricky issues pertaining to action individuation. For discussion see Hornsby, J. (1998), 'Action', in E. Craig (Ed.), *Routledge Encyclopaedia of Philosophy*. London: Routledge. Retrieved 04/02/08, from <http://www.rep.routledge.com/article/V001>

⁴⁹ I believe this view may also be informative in connection with other debates. In particular, I am interested to explore the relation between the claim that dependency on others is necessary for the autonomous performance of certain actions, and various formulations of republican

theories of freedom, which seem to require that the agent is not dependent upon, or dominated by, another agent. For formulations of these views, see Skinner, Q. (1998) *Liberty Before Liberalism*, Cambridge, Cambridge University Press; Pettit, P. (1997) *Republicanism: a theory of freedom and government*, Oxford, Oxford University Press; and for discussion, Dagger, R. (2005) 'Autonomy, Domination, and the Republican Challenge to Liberalism' in *Autonomy and the Challenges to Liberalism*, Christman, J. and Anderson, J. (eds.), Oxford, Oxford University Press, pp.117-203.

I anticipate that explicating the connections and tensions would be a large task. so set this aside for future research.

Conclusion

In this thesis I have explored the question of how autonomy might be relationally constituted. I argued against constitutively relational conceptions of autonomous choice; such conditions, I claimed, were unmotivated. I then argued against incorporating constitutively relational conditions into a conception of the autonomous agent. Such a conception is ill-placed to cohere with the normative framework of protecting from paternalism that accounts, in part, for autonomy's value. But in the final chapter, I showed how we could nonetheless go beyond the claim that social relations impacted on autonomy merely causally. The exercise of autonomy in a range of actions is constituted by social relations. I suggested a framework that focuses on the socially constituted authority of agents, as a way of fruitfully thinking about how social relations might hinder autonomous action, and set out the motivations for further pursuing this line of inquiry.

We started with the thought from Griffin that “Even if I constantly made a mess of my life, even if you could do better if you took charge, I would not let you do it. Autonomy has a value of its own”.¹ I hope to have, in the course of the thesis, clarified the ways in which autonomy is valuable. Autonomy is valuable in part due to the normative benefits to which it entitles agents. Moreover, if the suggestion I have made in the third part of the thesis can be further substantiated, it will be clear that autonomy has value because of its connection to the social relations in which agents stand. Being recognised as an autonomous agent – as having a kind of authority – enables agents to engage in a range of actions with others; namely, communicative and cooperative actions. If this is so, we can further understand, now, Griffin's objection to having his life taken over. For if an agent is not treated as an autonomous agent, and is not treated as an authority, ‘in charge’ of her life, her autonomy will be significantly curtailed. There will simply be a whole range of things that she can no longer do; namely, many of those things that we do with the help of, and in concert with, others.

¹ Griffin, J. (1986) *Op.Cit*, p. 67

Bibliographic references

Anderson, E. (2002) 'Should Feminists Reject Rational Choice Theory?' in *A Mind of One's Own*, Witt, C. and Antony, L. (eds.) Westview Press, pp.369-387

Antony, L. (1995) 'Is Psychological Individualism a piece of Ideology?' *Hypatia* 10 (3), pp.157-173

Arpaly, N. (2003) 'Varieties of Autonomy' in her *Unprincipled Virtue*, Oxford University Press, pp.117-149

Arpaly, N. and Schroeder, T. (1999) 'Praise, Blame and the Whole Self' *Philosophical Studies* 93 (2), pp.161-188

Arneson, R. (1980) 'Mill Versus Paternalism', *Ethics* 90 (4), pp.470-489

Austin, J.L. (1962) *How To Do Things With Words*, Cambridge, Mass. Routledge

Barclay, L. (2000) 'Autonomy and the Social Self' in MacKenzie and Stoljar (2000), pp.52-71

Baron, M. (1985) 'Servility, Critical Deference and the Deferential Wife' *Philosophical Studies* 48 (3), pp.393-400

Benhabib, S. (1994) 'Deliberative Rationality and Models of Democratic Legitimacy' *Constellations* 1 (1), pp.26-52

Benn, S.I. (1982) 'Individuality, Autonomy and Community' *Community as a Social Ideal*, Kamenka, E. (ed.) New York, St Martin's Press, pp.43-62

Benson, P. (2005a) 'Taking Ownership: Authority and Voice in Autonomous Agency' in Christman, J. and Anderson, J. (eds.) *Autonomy and the Challenges to Liberalism* 2005, Oxford, Oxford University Press, pp.101-126

Benson, P. (2005b) 'Feminist Intuitions and the Normative Substance of Autonomy' in Taylor, J.S. (ed.) *Personal Autonomy*, Cambridge, Cambridge University Press, pp.124-142

Benson, P. (2004) 'Blame, Oppression and Diminished Moral Competence', in *Moral Psychology: Feminist Ethics and Social Theory*, DesAutels, P. & Walker, M.U. (eds.) Rowman & Littlefield, pp.183-200

Benson, P. (2000) 'Feeling Crazy: Self-worth and the Social Character of Responsibility' in *Relational Autonomy*, MacKenzie and Stoljar (eds.) pp.72-94

Benson, P. (1994) 'Free Agency and Self-Worth' *The Journal of Philosophy* 91 (12) pp.650-668

Benson, P. (1991) 'Autonomy and Oppressive Socialisation' *Social Theory and Practice* 17, pp.385-408

Berlin, I. (1969) 'Two Concepts of Liberty' in his *Four Essays on Liberty*, Oxford, Oxford University Press

Bird, A. (2002) 'Illocutionary silencing' *Pacific Philosophical Quarterly* 83, pp.1-15

The Birth Trauma Association, accessed on 07/06/08

<http://www.birthtraumaassociation.org.uk/index.htm>

Bratman, M. 'Intention, Belief, Practical, Theoretical' (ms.) forthcoming in Timmerman, J. Skorupski, J. and Robertson, S. *Spheres of Reason*, Oxford University Press.

Bratman, M. (2005) 'Planning Agency, Autonomous agency' in Taylor J.S. (ed.) *Personal Autonomy* Cambridge, Cambridge University Press, pp.33-57

Bratman, M. (1992) 'Shared Cooperative Activity' *The Philosophical Review* 101 (2), pp.327-341

- Bordo, S. (2003) 'Are Mothers Persons?' in her *Unbearable Weight: Feminism, Western Culture and the body*, California, University of Berkeley Press, pp.71-97
- Buss, S. (2005) 'Valuing Autonomy and Respecting Persons: Manipulation, Seduction and the Basis of Moral Constraint' *Ethics* 115 pp.195-235
- Card, C. (1995) 'Gender and Moral Luck' in Held, V. (ed.) *Justice and Care*, Westview Press, pp.70-100
- Chafe, W.H., Gavins, R., Korstad, R. (eds.) (2001), *Remembering Jim Crow*, New York, The New Press
- Chambers, C. (2008) *Sex, Justice and Culture: The Limits of Choice*, Pennsylvania, Pennsylvania University Press
- Christman, J. (2006) Book Review: *Autonomy Gender, Politics* Friedman, M. (2003), *The Journal of Value Inquiry* 40, pp.137-143
- Christman, J. (2005) 'Autonomy, Self-Knowledge and Liberal Legitimacy' in Christman, J. and Anderson, J. (eds.) *Autonomy and the Challenges to Liberalism* 2005, Oxford, Oxford University Press, pp.330-359
- Christman, J. (2004) 'Relational Autonomy, Liberal Individualism, and the Social constitution of Selves' *Philosophical Studies* 117, pp.143-164
- Christman, J. (1991a) 'Liberalism and Individual Positive Freedom' *Ethics* 101 (2), pp.343-359
- Christman, J. (1991b) 'Autonomy and Personal History' *Canadian Journal of Philosophy* 21, pp.1-24
- Code, L. (1991) "Second Persons" in her *What Can She Know? Feminist Theory and the Social Construction of Knowledge*, New York, Cornell University Press
- Connor, S. (2006) 'Scientist's Cloning claims Untrue ... Except for the Dog' *Independent on Sunday* Jan, 11 2006, accessed 11/05/08

<http://www.independent.co.uk/news/science/scientists-cloning-claims-untrue-except-for-the-dog-522484.html>

Dagger, R. (2005) 'Autonomy, Domination, and the Republican Challenge to Liberalism' in Christman, J. and Anderson, J. (eds.) *Autonomy and the Challenges to Liberalism*, 2005, Oxford, Oxford University Press, pp.117-203

Darwall, S. (1977) 'Two Kinds of Respect' *Ethics* 88 (1), pp.36-49

Davidson, D. (1984) 'Communication and Convention' in his *Inquiries into Truth and Communication*, Oxford, Oxford University Press

Davidson, D. (1980) *Essays on Actions and Events*, Oxford: Oxford University Press

Dworkin, G. (1988) *The Theory and Practice of Autonomy*, Cambridge, Cambridge University Press

Dworkin, G. (1983) 'Paternalism: Some Second Thoughts' Sartorius, R. (ed.) *Paternalism*. Minneapolis, University of Minnesota Press, pp.105-112

Dworkin, G. (1972) 'Paternalism' *Monist* 56, pp.64 – 8

Feinberg, J. (1986) *Harm to Self: The Moral Limits of the Criminal Law, Volume Three*, Oxford, Oxford University Press

Fischer, J. & Ravizza, M. (1998) *Responsibility and Control: A Theory of Moral Responsibility*, New York, Cambridge University Press

Frankfurt, H. (1999) 'The Faintest Passion' in his *Necessity Volition and Love* Cambridge, Cambridge University Press, pp.95-107

Frankfurt, H. (1971) 'Freedom of the Will and the Concept of a Person' *The Journal of Philosophy* 68 (1), pp.5-20

Frankfurt, H. (1969) 'Alternate Possibilities and Moral Responsibility'. *The Journal of Philosophy* 66 (23), pp.829-839

Fricker, M. (2007) *Epistemic Injustice: Power and the Ethics of Knowing*, Oxford, Oxford University Press

Friedman, M (2003) *Autonomy, Gender, Politics* New York, Oxford University Press

Friedman, M. (1993) 'Beyond Caring: The Demoralisation of Gender' in Laraabee, M.J. (ed.) *An Ethic of Care: Feminist and Interdisciplinary Perspectives*, Routledge, pp.258-275

Friedman, M. (1985) 'Moral Integrity and the Deferential Wife' *Philosophical Studies* 47 (1), pp.141-150

Gaus, G. (2005) 'The Place of Autonomy within Liberalism' in Christman, J. and Anderson, J. (eds.) *Autonomy and the Challenges to Liberalism* 2005, Oxford, Oxford University Press, pp.272-306

Gauthier, D. (1986) *Morals by Agreement*, Oxford, Oxford University Press

Gilbert, M. (1990) 'Walking Together: A Paradigmatic Social Phenomenon' *Midwest Studies in Philosophy* 15, pp.1-14

Gilbert, M. (1992) *On Social Facts*, Princeton, Princeton University Press

Gilligan, C. (1994) *In a Different Voice: Psychological Theory and Women's Development*, Harvard University Press

Griffin, J. (1988) *Well-being: Its Meaning, Measurement and Moral Importance*, Oxford, Oxford University Press

Hanrahan, R. and Antony L, (2005) 'Because I said so: towards a feminist theory of authority' *Hypatia* 20 (4), pp.59-79

Harris, A. (1993) 'Race and Essentialism in Feminist Legal theory' in Weisberg, D. K. (ed.) *Feminist Legal Theory: Foundations*, Philadelphia: Temple University Press, pp.349-358

Held, V. (1993) *Feminist Morality*, Chicago, University of Chicago Press

Hill Jr., T. (1991a) 'Servility and Self-Respect' in his *Autonomy and Self Respect*, Cambridge, Cambridge University Press, pp.4-18

Hill Jr. T. (1991b) 'Self-respect Reconsidered' in his *Autonomy and Self-Respect*, Cambridge, Cambridge University Press, pp.19-23

Hill, Jr. T, (1991c) 'The Importance of Autonomy' in his *Autonomy and Self-Respect*, Cambridge, Cambridge University Press, pp.43-51

Holroyd, J. 'The Metaphysics of Relational Autonomy' (forthcoming) *Feminist Metaphysics* Witt, C. (ed.) Springer Publishing

Hornsby, J. (2004) 'Agency and Actions' Steward, H. and Hyman, J. (eds.) *Agency and Action*, Cambridge, Cambridge University Press, pp.1-23

Hornsby, J. (1998), 'Action', in E. Craig (Ed.), *Routledge Encyclopaedia of Philosophy*. London: Routledge. Retrieved 04/02/08, from <http://www.rep.routledge.com/article/V001>

Hume, D. *A Treatise of Human Nature* (1888/1978) Nidditch, P.H. (ed.), 2nd Edition (with Analytical Index by L.A. Selby-Bigge), Oxford, Clarendon Press

Hursthouse, R. (1999) *On Virtue Ethics*, Oxford, Oxford University Press.

Jacobson, D. (1995) 'Freedom of Speech Acts? A response to Langton' *Philosophy and Public Affairs* 23(1), pp.64-79

Kant, I. *The Metaphysics of Morals*, (1797/1996) Gregor, M. (ed.), Cambridge, Cambridge University Press

Klarman, M.J. (2004) *From Jim Crow to Civil Rights: the Supreme Court and the Struggle for Equality*, Oxford, Oxford University Press

Kohlberg, L. (1984) *The Psychology of Moral Development: The Nature and Validity of Moral Stages*, Harper Collins

Korsgaard, C. (2002) 'Autonomy, Efficacy and Agency' *Locke Lecture III*. Accessed 04/05/08 at:

<http://www.people.fas.harvard.edu/~korsgaard/#Locke%20Lectures>

Korsgaard, C. (2002) 'Practical Reason and the Unity of the Will' *Locke Lecture II*. Accessed 04/05/08 at:

<http://www.people.fas.harvard.edu/~korsgaard/#Locke%20Lectures>

Korsgaard, C. (1996a) *The Sources of Normativity* O'Neill, O. (ed.), Cambridge, Cambridge University Press

Korsgaard, C. (1996b) 'Kant's Formula of Humanity' in her *Creating the Kingdom of Ends*, Cambridge, Cambridge University Press, pp.106-132

Kutz, C. (2000) 'Acting Together' *Philosophy and Phenomenological Research* 61 (1) pp.1-31

Langton, R. (1993) 'Speech Acts and Unspeakable Acts' *Philosophy and Public Affairs* 22 (4), pp.293-330

Langton, R. and Hornsby, J. (1998) 'Free Speech and Illocution' *Journal of Legal Theory* 4, pp.21-37

Langton, R. and West, C. (1999), 'Scorekeeping in a Pornographic Language Game' *Australasian Journal of Philosophy* 77 (3), pp.303-319

Lenman, J. (2000) 'Consequentialism and Cluelessness' *Philosophy and Public Affairs* 29 (4), pp.342-370

Lessing, D. (1950) *The Grass is Singing* London, Harper Perennial

Lewis T.M., McClain M., Pittenger A. (1997) 'Alien Hand Syndrome and Sensory Ataxia: A Case Study of a Unique Presentation of the Disorder', *Archives of Clinical Neuropsychology* 12 (4), pp.357-358

Luker, K. (1975) *Taking Chances: abortion and the decision to contracept*, Berkeley, University of California Press.

MacIntyre, A. (1973) 'The Essential Contestability of Some Social Concepts' *Ethics* 84 (1), pp.1-9

MacKinnon, C. (1987) *Feminism Unmodified*, Harvard University Press.

MacKenzie, C. and Stoljar, N. (2000) 'Autonomy Re-figured' in their *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Self*, Oxford, Oxford University Press, pp.3-34

Maitra, I. (2004) 'Silence and Responsibility', *Philosophical Perspectives* 18 (1), pp.189-208

Manson, N.C. and O'Neill, O. (2007) *Rethinking Informed Consent in Bioethics*, Cambridge, Cambridge University Press

McKenna, M. (2005) 'The Relationship between Autonomous and Morally Responsible Agency', in Taylor, J.D. (ed.), *Personal Autonomy*, Cambridge, Cambridge University Press, pp.205-235

Mele, A. (2005) 'Agnostic Autonomism Revisited' Taylor (ed.) *Personal Autonomy*, pp.109-124

Mele, A. (1995) *Autonomous Agents: from Self-Control to Autonomy*, Oxford, Oxford University Press

Mendus, S. (1999) 'Out of the Doll's House: Reflections on Autonomy and Political Philosophy' *Philosophical Explorations* 2 (1), pp.59-70

- Mendus, S. (1992) Review: *Autonomy and Self-Respect*, Hill Jnr., T. (1991) *Philosophy* 67 (262), pp.561-563.
- Meyers, D. (2002) *Gender in the Mirror: Cultural Imagery and Women's Agency* Oxford, Oxford University Press
- Meyers, D. (2000) 'Intersectional Identity and the Authentic Self?: Opposites Attract!' *Relational Autonomy* MacKenzie and Stoljar, (eds.) pp.151 – 181
- Meyers, D (1987) 'Autonomy and the Paradox of Feminine Socialization' *The Journal of Philosophy* 84 (11), pp. 619-628
- Mill, J.S. Utilitarianism, Warnock, M. (ed.), *Mill: Utilitarianism and Other Writings*, Collins, Glasgow, 1962
- Miller, S. (1992) 'Joint Action' *Philosophical Papers* 21 (3) pp.275-297),
- Narayan, U. (2002) 'Minds of Their Own: Choices, Autonomy, Cultural Practices and Other Women' in Antony, M. and Witt, C. (eds.), *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, Boulder, Colorado: Westview Press, pp. 418-432
- Narayan, U. (1997) *Dislocating Cultures: Identities, Traditions and Third World Feminism*, New York, Routledge
- O'Neill, O. (2003) 'Autonomy: The Emperor's New Clothes' *Aristotelian Society Supplementary Volume* 77 (1), pp.1-21
- Okin, S. M. (1989) *Justice, Gender and the Family*, NY, Basic books
- Oshana, M. (2006) *Personal Autonomy in Society*, Hampshire, Ashgate Publishing Ltd
- Oshana, M. (2005) 'Autonomy and Free Agency' in Taylor, J.S. (ed.), *Personal Autonomy*, Cambridge, Cambridge University Press
- Oshana, M. (2003) 'How much should we value autonomy?' in Frankel Paul, E., Miller Jnr., F.D., Paul, J. (eds.) *Autonomy* Cambridge University Press, pp.99-126

Oshana, M. (1998), 'Personal Autonomy and Society' *The Journal of Social Philosophy* 29 (1), pp. 81-102

Parfit, D. (2002) 'What we Could Rationally Will: I. Rational Consent', *The Tanner Lectures on Human Value*. pp.287-369

<http://www.tannerlectures.utah.edu/lectures/atoz.html#p>

Pettit, P. (1997) *Republicanism: a Theory of Freedom and Government*, Oxford, Oxford University Press.

Pritchard, D. (2005) *Epistemic Luck*, Oxford, Oxford University Press

Rawls, J. (1971) *A Theory of Justice*, Cambridge, Harvard University Press

Rawls, J. (1993) *Political Liberalism*, Columbia University Press

Raz, J. (1986) *The Morality of Freedom*, Oxford University Press

Reed, A. 'Not a Happy Birthday' *The F word: Contemporary UK Feminism*, http://www.thefword.org.uk/features/2008/03/not_a_happy_bir accessed 07/06/08

Saul, J. (2003) *Feminism: Issues and Arguments*, Oxford, Oxford University Press

Scanlon, T.M. (1998) *What We Owe to Each Other*, Mass. Harvard University Press

Scanlon, T.M. (1986) 'The Significance of Choice' *The Tanner Lectures on Human Values*, delivered Brasenose College, Oxford, May 1986.

Scoccia, D. (2008) 'In Defense of Hard Paternalism' *Law and Philosophy* 27, pp.351-381.

Scoccia, D. (1990) 'Paternalism and Respect for Autonomy' *Ethics* 100 (2), pp.318-334

Skinner, Q. (1998) *Liberty Before Liberalism*, Cambridge, Cambridge University Press

- Slote, M. (1980) 'Understanding Free Will' *The Journal of Philosophy* 77 (3) pp.136-151
- Stoljar, N. (2000) 'Autonomy and the Feminist Intuition' in MacKenzie, C. and Stoljar, N. (eds.) *Relational Autonomy: Feminist Perspectives on Autonomy and Agency*, Oxford, Oxford University Press, pp.94-111
- Superson, A. (2005) 'Deformed Desires and the Informed Desire Test' *Hypatia* 20 (4), pp.109-126
- Swimm, D. 'More than a Traumatic Birth' *TrueBirth.com*, accessed 06/06/08
<http://www.truebirth.com/2008/02/10/more-than-a-traumatic-birth/>
- Taylor, J.S. (2003) 'Autonomy, Duress and Coercion', *Social Philosophy and Policy* 20 (2), pp.127-155
- Tuomela, R. (2005) 'We-intentions Revisited' *Philosophical Studies* 135 (3), pp.327-369
- Tuomela, R. and Miller, K. (1988) 'We Intentions' *Philosophical Studies* 53, pp.367-398
- Velleman, D. (1997) 'How to Share an Intention' *Philosophy and Phenomenological Research* 57 (1), pp.29-50
- Waldfoegel, J. (2007) 'TV Is Good for You, if you are a woman in rural India, at least.'
Slate.com Posted Monday, Aug. 20, 2007, accessed 08/05/08
<http://www.slate.com/id/2172474/>
- Watson, G. (1975) 'Free Agency' *The Journal of Philosophy* 72 (8), pp.205-220
- Westlund, A. (2003) 'Selflessness and Responsibility for Self' *The Journal of Philosophy* 112 (4), pp.483-523
- Westlund, A. 'Women's autonomy in conditions of oppression' (unpubl. ms.)

Westlund, A. 'Rethinking Relational Autonomy' (forthcoming, *Hypatia* 2009)

Wieland, N. (2007) 'Linguistic Authority and Convention in a Speech Act Analysis of Pornography' *Australasian Journal of Philosophy* 85 (2), pp.435-456

Wolf, S. (1980) 'Asymmetrical Freedom' *The Journal of Philosophy* 77 (3), pp.151-166

Woolf, V. (1931) 'Professions for Women' a speech before the National Society for Women's Service, in Woolf, V. *The Virginia Woolf Reader*, Leaska, M. (ed.). (1984).