

Feminism, Liberalism, and Relational Autonomy

By

Emily McGill-Rutherford

Dissertation

Submitted to the Faculty of the  
Graduate School of Vanderbilt University

In partial fulfillment of the requirement

For the degree of

DOCTOR OF PHILOSOPHY

In

Philosophy

August, 2015

Nashville, Tennessee

Committee:

Marilyn Friedman (Chair)

Larry May

Robert Talisse

Diana Tietjens Meyers

## TABLE OF CONTENTS

Introduction.....	i
I: The Unsteady Alliance of Liberalism and Feminism: Political Liberalism and the Individualism Critique.....	1
1. Political Liberalism.....	2
1.1 Political vs. Comprehensive Liberalism.....	2
1.2 Neutrality.....	4
2. Anti-Liberal Criticism.....	7
2.1 The Ontological Objection.....	8
2.2 The Epistemic Objection.....	13
2.3 The Identity Objection.....	14
3. A Political Liberal Response.....	17
3.1 Response to the Ontological Objection.....	18
3.2 Response to the Epistemic Objection.....	23
3.3 Response to the Identity Objection.....	27
4. Conclusion.....	29
II: Traditional Autonomy: Hierarchical, Historical, and Normative Competency Theories.....	31
1. Distinctions and Classifications.....	31
2. Hierarchical Theories.....	32
2.1 Harry Frankfurt.....	33
2.2 The Regress/Arbitrariness Problem and Wholeheartedness.....	34
2.3 The Vagueness of Wholehearted Identification.....	36
2.4 The Manipulation Problem.....	37
2.5 The Objectivity Problem.....	38
3. Historical Theories.....	40
3.1 John Christman.....	40
3.2 The Agentic Authority Problem.....	43
3.3 The Limited Subjectivity Problem.....	44
4. Normative Competency Theories.....	45

4.1 Gary Watson.....	45
4.2 The Influence of Desire Problem.....	48
III: Recent Work on Relational Autonomy.....	50
1. Relational Autonomy: The Early Years.....	50
1.1 The Self: United, Transparent, Rational, Individual?.....	50
1.2 Autonomy: Isolationist, Individualistic, Masculine?.....	54
1.3 Reclaiming Autonomy.....	59
2. Procedural or Substantive: Never the ‘Twain Shall Meet.....	60
2.1 From Procedural to Substantive.....	60
2.2 Autonomy as Social-Relational.....	63
3. Breaking Dichotomies.....	66
3.1 Autonomy as Dialogical.....	67
3.2 Utilizing Value.....	68
3.3 Mainstreaming Feminism.....	70
IV: Neither Islands Nor Dupes: Relational Autonomy and the Self.....	73
1. Why Relational?.....	75
2. Socially Constituted Selves.....	75
2.1 Social Causation.....	76
2.2 Social Constitution.....	79
2.3 Social Determination.....	81
3. Concerns with the Social Constitution of Selves.....	83
3.1 Questioning Social Constitution.....	83
3.2 What Are Social Relationships?.....	85
4. Socially Constituted Autonomy.....	86
4.1 Social Autonomy: Two Interpretations.....	87
5. Concerns with the Social Constitution of Autonomy.....	89
5.1 Deformed Desires.....	89
6. A Closer Look at Feminist Strong Substantive Autonomy.....	91
6.1 Three Strong Substantive Accounts.....	91
6.2 Deformed Desires, Revisited.....	94
7. Relational, Reconsidered.....	96

V: Purely Procedural Feminist Autonomy.....	98
1. What We Want and What it Costs.....	98
1.1 What Do We Want?.....	98
1.2 What Does it Cost?.....	101
1.3 Feminist Husbands and Dominant Wives.....	103
2. Agent Centered, Liberty Oriented.....	105
2.1 Reflective Endorsement.....	106
2.2 Diachronic Value Cohesion.....	109
2.3 Narratability.....	112
3. A Hard Case: There is No ‘You’ in Quivering.....	116
4. Biting Bullets?.....	123
Bibliography.....	130

## INTRODUCTION

“[Women] are so far in a position different from all other subject classes, that their masters require something more from them than actual service. Men do not want solely the obedience of women, they want their sentiments. All men, except the most brutish, desire to have, in the woman most nearly connected with them, *not a forced slave but a willing one*, not a slave merely but a favourite. They have therefore put everything in practice to *enslave their minds* [...] All women are brought up from the very earliest years in the belief that their ideal of character is the very opposite of that of men; not self will, and government by self-control, but *submission, and yielding to the control of others*.” – John Stuart Mill, “The Subjection of Women”

In 1869, liberal theorist John Stuart Mill spoke of the devastating effects of women’s oppression. In a patriarchal society, when a woman’s wellbeing (and indeed, survival) depends upon earning and keeping the favor of a man, it becomes necessary for her to bend her will to meet the expectations of the man in power. Women are taught, by other women, not only to *do* what pleases men but also to *desire* it. We have been taught to be slaves for such a long time that we have become willing ones; we have internalized – and become participants in – our own oppression.

In Shakespeare’s *The Merchant of Venice*, Shylock exclaims: “The villainy you teach me I will execute – and it shall go hard but I will better the instruction.” It seems women, too, have bettered our instruction. The gender wage gap, the epidemic of sexual assaults on college campuses, and the sexual enslavement of girls and women have all been criticized as exaggerations spread by feminist mobs – and those leveling the criticisms have been women.<sup>1</sup> A movement of “anti-feminist women” has yielded an internet activist campaign in which a vast array of young women hold signs proclaiming why *they* don’t need feminism. “I don’t need feminism because I’m not a manipulative idiot playing victim!” one reads. Another young woman says that she doesn’t need feminism, because real feminism isn’t about “abortions, free

---

<sup>1</sup> See, for example, the work of Christina Hoff Sommers, including a recent article in TIME, “5 Feminist Myths That Will Not Die” (September 2, 2014). <http://time.com/3222543/5-feminist-myths-that-will-not-die/>

birth control, and the ability to walk around like a shameless slut while damning the male population!” A third holds a sign that reads, “I don’t need feminism because *those rape statistics are made up anyway shut up rape victims!!!*” And a fourth confesses, “I don’t need feminism because I want boys to like me.”

These sentiments are especially heartbreaking coming from young women and teenage girls. Perhaps we’ve come to expect people like Pat Robertson arguing, “the feminist agenda is not about equal rights for women. It is about a socialist, anti-feminist political movement that encourages women to leave their husbands, kill their children, practice witchcraft, destroy capitalism and become lesbians.”<sup>2</sup> But it’s not so easy to roll our eyes when the same thoughts come from our sisters, friends, and daughters. Is this what Mill had in mind when he talked about enslaving the minds of women? Is this what internalized oppression looks like?

More recently, feminist theorists have questioned whether or not contemporary liberal theorists can take account of internalized oppression at all. The liberal focus on individual choice, it is argued, obscures the way that a patriarchal society can influence women’s desires.

Linda Zerilli has recently argued, for example:

Liberal feminists [...] more or less take for granted that what holds women back is not what women themselves desire but what men have put in women’s way whenever they seek to escape the confines of the traditional family [...] But what if the truly pernicious effect of male power turns out to be that stereotypical images of women are ‘most deeply injurious at the point at which they are empirically real,’ as MacKinnon puts it [...] If we accept the notion that individuals are more or less potent and potential makers of meaning who face obstacles that need to be removed, then we may well fail to see the very nature of the constraint that laws, custom, and rules of a male-dominated society place on women. We will fail to see that many women have become what a male-centered society has allowed them to be. Consequently, their conventionally feminine choices will appear to be free (Zerilli 2015: 374-5).

---

<sup>2</sup> “Equal Rights Initiative in Iowa Attacked,” Washington Post, August 1992.

This political critique crosses over into the feminist literature on autonomy. Traditional theories of autonomy, like those of Harry Frankfurt and Gerald Dworkin, are thought to fail because they accept *liberal individualism*. It is as if traditional autonomy theorists and contemporary liberals imagine autonomy as the absence of chains; insofar as women are chained, they are non-autonomous. But if there are no chains, on the other hand, women are autonomous. The feminist argument is that women's oppression doesn't function like chains; it functions instead like a computer virus, sneaking into our minds when we aren't paying attention and rerouting our beliefs. If liberal individualism prevents traditional autonomy theorists from understanding that this rerouting, too, can compromise our autonomy, then so much the worse for traditional autonomy.

The primary purpose of this project is to respond to the feminist critique of traditional theories of autonomy, which revolves around the charge that such theories are *too individualistic*. Feminists like Natalie Stoljar and Jennifer Nedelsky, for example, argue against the liberal atomism that they see at the center of traditional autonomy theories. Their resulting theory of relational autonomy is meant to remedy the following two shortcomings: first, that traditional theories of autonomy posit an individualistic conception of the self, and second, that such theories posit an individualistic conception of autonomy. Instead, feminists have argued for a theory of autonomy that takes account of the ways in which persons are irreducibly social, and the ways in which autonomy itself is only possible within certain types of social relationships. Relational accounts are thus meant to be more *descriptively* accurate than traditional theories, but they are also meant to be more *normatively* sound. Not only do such accounts better describe the lives of women in oppressive societies (where traditional accounts fail due to their focus on isolated individuals). Relational accounts also vindicate the normative commitments of

feminism; specifically, these accounts are better able to conceptualize and work to eliminate the oppression of women.

The project proceeds in the following way: first, I separate the feminist charges against liberalism from the feminist charges against traditional autonomy. Although these two issues run together in much of the literature, it is at least *prima facie* possible to consider personal autonomy separately from political autonomy. I thus isolate the feminist critique of *liberal atomism* from the feminist critique of *individualistic autonomy*. The first chapter examines the feminist critique of liberalism to determine why it is that the charge of liberal atomism continues to stick, even though it has been clearly established by both feminist liberals – like Martha Nussbaum and Ruth Abbey – and mainstream liberals – like John Rawls and Will Kymlicka – that liberalism is not guilty of such a charge. While this chapter does not provide a full defense of liberalism, it does argue that such a defense is possible, and possible in a way that upholds feminist goals.

Having separated the critiques of liberalism from the critiques of traditional autonomy, the rest of the dissertation focuses exclusively on the autonomy debate. Chapter Two provides an overview of the traditional debate thus far, and Chapter Three provides an overview of the feminist responses to this debate. In addition to reviewing the history of the feminist debate, Chapter Three discusses the most recent feminist literature on autonomy, including the newly released *Autonomy, Oppression, and Gender* (2014). While the debate was young for the release of MacKenzie and Stoljar's groundbreaking *Relational Autonomy* (2000), it has since matured. My third chapter takes account of this maturation and the changes that have occurred in the debate – something which has not yet been offered elsewhere.



Chapter Four contains an argument against the most robustly relational feminist accounts of autonomy – those which accept a relational account of both the self and autonomy. I argue that, although such theories are explicitly designed to vindicate the normative commitments of feminism, their implications yield unwelcome results from this very same feminist perspective. This chapter attempts to switch the burden of proof from thinly relational theories (like those offered by Marilyn Friedman, John Christman, and myself) to robustly relational theories. The goal is for Chapter Four to thus pave the way for the final chapter, Chapter Five, in which I present my own *purely procedural* account of autonomy. My account has the following features: 1. It is purely procedural, that is, it does not place direct normative constraints on an agent's preferences; 2. It is relational in that it explicitly acknowledges the influence of social factors on both selfhood and autonomy, but 3. It is individualistic in that it does not allow social factors to play a constitutive role in selfhood or agency; and 4. It is able to address the problem of internalized oppression and adaptive preferences – what I take to be the central issue in the current feminist debate.

Because my account is not constitutively social, it is unlikely to be characterized as a *relational* account at all. I argue, however, that it is this very feature which allows my account to vindicate the feminist normative commitments espoused by constitutively relational accounts. This sets my theory apart from the *vast* majority of feminist accounts of autonomy. It is my hope to shift the feminist debate away from robustly relational theories and back toward a thin sort of individualism. And although it will certainly be charged that this situates my theory squarely outside of the feminist debate, I maintain that my account is motivated by the very same normative commitments that moved feminists toward relational autonomy in the first place.

## CHAPTER I

### THE UNSTEADY ALLIANCE OF LIBERALISM AND FEMINISM: POLITICAL LIBERALISM AND THE INDIVIDUALISM CRITIQUE

Much contemporary feminist philosophy shares a common foe outside of those enemies we might expect (patriarchy, gender inequality, and women's oppression). This common foe is *liberalism* - a framework for political theorizing that is presupposed in the institutions of modern democracy. Feminist scholar Jennifer Nedelsky explains that feminists often feel "rage at all things liberal" (Nedelsky 1989: 9), and Claudia Card notes that liberalism has become the new "L-word" (Card 2006). Non-liberal feminists, she notes, approach liberalism with an attitude that is "scornful, condescending, caricaturing" (Card 2006: 223). In the minds of some feminists, Card recognizes, liberalism has been "something to be cured of, like a disease, or a wrong turn to be repented" (Card 2006: 224).

But why repent? Where has liberalism gone so wrong? My aim in this chapter is to show that one type of liberalism - *political* liberalism - properly understood, is no foe to feminism; it is instead a powerful ally. I will do so by focusing on one subset of feminist criticisms, those focused on the (allegedly) individualist nature of liberal theory.<sup>1</sup> I group these criticisms together as the *Individualism Critique*, which can be summarized in the following way:

---

<sup>1</sup> The fact that I am focusing on this single subset of criticisms should not be taken to imply that it is the only type of feminist criticism, or that it is the most important. The criticisms discussed in this chapter are united by a focus on individualism and neutrality, but there are many other valuable anti-liberal feminist insights that I don't have room to consider here. These include debates about (for example) the public/private distinction, the role of family and community in the moral upbringing of citizens, the standing of children within liberal societies, and the role of civic education. For excellent collections of recent discussions see Baehr 2004, Abbey 2011, and Abbey 2013.

**Individualism Critique: Liberalism rests on a fundamentally flawed conception of selves as individuals, which in turn requires an incoherent neutralism on the part of the state. Liberalism must therefore be rejected.**

I argue that there are at least two problems with the *Individualism Critique*: first, it relies on a fundamental misunderstanding of liberalism. Second, this misunderstanding causes liberal and non-liberal feminists to talk past one another, which in turn leads to entrenched disagreements and unproductive debates. Criticisms that were first raised nearly thirty years ago are still being made. Why? This chapter serves partly as a diagnostic, but also as a call for increased discussion across the divide.

## **1. Political Liberalism**

### ***1.1. Political vs. Comprehensive Liberalism***

First, I have to explain what political liberalism is. *Liberalism* itself is a motley of very different kinds of views. One taxonomical tool developed by John Rawls – the distinction between political and comprehensive liberalism - is helpful (Rawls 1985). Political liberalism is a political theory that does not require a specific moral, metaphysical, or epistemological doctrine. Instead, it is meant to be neutral between different such theories. Thus, a citizen can hold any number of different moral theories while still being politically liberal. In Rawls' language, political liberalism is a "freestanding view" (Rawls 2005: 12). Comprehensive liberalism, on the other hand, is not. Instead, comprehensive liberalism *does* rely on a given moral, metaphysical, or epistemological foundation - it is not neutral between all such doctrines.

hand, is not. Instead, comprehensive liberalism *does* rely on a given moral, metaphysical, or epistemological foundation - it is not neutral between all such doctrines.

Speaking broadly, political liberalism rests upon a normative understanding of society as a system of cooperation among free and equal citizens. Part of what it means to respect citizens as free and equal is to require that any coercive state action be justifiable to the citizens of that state. This thought, that state coercion requires justification, is the liberal principle of legitimacy (Rawls 2005: 137; Talisse 2012: 27). Let me now stipulate that, for the remainder of this chapter, the sort of liberalism I defend is the political sort – not the comprehensive sort. To explain why, it is important to note that political liberalism was developed in order to rectify what Rawls saw as a misguided effort - the effort to justify liberalism by an appeal to a specific moral doctrine (Gaus 2003: 195; Talisse 2012: 28). Not only did Rawls think that this effort was bound to fail, but he also thought that such a justification would ultimately fail to respect all citizens as free and equal persons (Rawls 2005: 9-10; cf. Quong 2011: 2).<sup>2</sup> In modern societies, people hold a variety of different moral views or comprehensive doctrines. Any attempt to justify state action by appeal to any one of these moral views will ultimately leave out those who adopt a different perspective. Remember that the liberal principle of legitimacy requires that any coercive state action be justifiable to all citizens, regardless of which moral view they happen to adopt. Attempting to justify liberalism by appealing to one moral doctrine will fail to justify coercive state action to those citizens who have a different perspective - and *this is unacceptable*. In fact, it is *unjust*. While comprehensive liberals may ground their political theory in a given moral

---

<sup>2</sup> Rawls argues that any attempt to establish liberalism by appealing to a particular comprehensive doctrine is ultimately going to be self-undermining. Appealing to the fact of reasonable pluralism, Rawls notes that modern societies have many different citizens with different comprehensive doctrines. If we appeal to our own comprehensive doctrine to determine what the state should do, we will end up forcing this doctrine on others.

to justify the coercive use of state power to all citizens, regardless of their comprehensive doctrine. The respect for all citizens as free and equal persons, then, is maintained.

This brief portrait has obviously presented an extremely limited account of one type of liberalism. There are many different ways to divide up different sorts of liberalisms, and I've barely scratched the surface. The important thing to note about political liberalism for now is that it is a theory about the relationship between the state and its citizens. Although comprehensive liberals might adopt a robust moral theory, this robust moral theory is not mandated by liberalism. Even comprehensive liberals are fundamentally concerned with coercive state actions - although they may ultimately disagree with political liberals about how such coercion can be justified. In other words, political liberalism is an articulation of how states should treat citizens; it is not a theory espousing the proper way for people to live in all aspects of their lives. In fact, political liberalism is committed to avoiding such commitments.

### ***1.2. Neutrality***

Liberalism is committed to neutrality - that is, to the commitment to avoid the endorsement of a specific moral theory. Because this concept will be central to my discussion, and because it is the source of so much conflict, I will take a moment to provide a brief account of liberal neutrality. Neutrality is entailed by certain commitments of liberalism, including the commitment to viewing citizens as free and equal persons.<sup>3</sup> There are, broadly speaking, three

---

<sup>3</sup> Having said this, it is not the case that every liberal endorses neutrality to the same extent. Joseph Raz, for example, denies that neutrality is required by liberalism. See Raz 1986. Perfectionists in general think that the state exists to make people good, where good is some positive concept (such as, for example, worshipping the right God or flourishing). There are two points to note here. First, I am not focusing on perfectionist liberalism in this chapter but am instead focusing on the sort of *political* liberalism motivated by Rawls's insight that any attempt to establish liberalism based on a particular comprehensive doctrine is going to be self-

types of neutrality: neutrality of aim or intent, neutrality of effect, and neutrality of justification (see e.g. Arneson 2003). Neutrality of aim requires that no state action should intend to promote one conception of the good over others; citizens within liberal societies have many varied conceptions of the good, and the state must not aim to treat people differently based on these different conceptions. Neutrality of aim, however, is reductive. Depending on the particular understanding of neutrality of aim, the idea reduces to either neutrality of effect or neutrality of justification (Patten 2012: 255-6).<sup>4</sup> Neutrality of effect requires that no state action should bring about the consequence that any particular conception of the good is advantaged over others (and this same idea holds for the adherents of these conceptions). Finally, neutrality of justification requires that all state action must be justified without reference to any particular conception of the good.

Because neutrality of aim reduces to either neutrality of effect or neutrality of justification, I will focus on the latter two. Neutrality of effect is often taken, by critics, to be the standard liberal view. Some feminist critics of liberalism charge the theory with preventing citizens from discussing controversial issues just because the issues happen to involve robust moral claims (Zerilli 2012: 12), and forcing the state to be silent about gender discrimination (Levey 2005). Carole Pateman eloquently (but mistakenly) charges that liberalism is “cut off in a neutral castle behind a moat without a drawbridge from ideas, beliefs and values that are sources of conflict” (Pateman 2002: 39). But one would be hard pressed to find a single contemporary liberal supporting neutrality of effect; from the liberal perspective, it is “unrealistic and undesirable” to require the state to be neutral in this way (Quong 2011: 18). Quite simply,

---

undermining (because you must ultimately appeal to reasons that other citizens don't recognize *as reasons*). For an excellent discussion of politically liberal objections to liberal perfectionism, see Quong 2011: 12-136.

<sup>4</sup> Thanks also to Andrew Forcehimes for help clarifying this point.

this is because almost any policy we can imagine is going to have non-neutral effects (Patten 2012: 256-7), and so the state would be forbidden to act in almost every situation. Rawls himself argues that it is impossible for a state “not to have important effects and influences as to which comprehensive doctrines endure and gain adherents over time; and it is futile to try to counteract these effects” (Rawls 2005: 193). It is no wonder that anti-liberal feminists have found so much wrong with liberal neutrality, when the concept is understood as neutrality of effect!

I take it that neutrality of justification is now the standard liberal view. Again, a state violates justificatory neutrality when it justifies a particular policy by appealing to a specific conception of the good (or set of conceptions). For example, if the state were to justify a federal maternity leave policy by appealing to the ‘fact’ that women are (biologically) designed to stay at home with their children, such a justification would fail the relevant type of neutrality. If the state were to justify the very same policy by appealing to the ‘fact’ that women are mandated by God to serve their children and their husbands, the policy would similarly fail. Why? Because in each case, the state is justifying a policy by appealing to a specific comprehensive doctrine, and this is unjust (recall our discussion in the previous section).

I will mention one more crucial point about justificatory neutrality. Some feminist critics charge that neutrality forces a state into inaction; I have demonstrated that this charge rests on a faulty understanding of liberal neutrality as neutrality of effect. Justificatory neutrality is not open to this same critique. If there is a consensus of reasonable comprehensive doctrines, the state can act accordingly. So, if there is a consensus that sexism is an unacceptable form of oppression (for example) the state can justifiably – and neutrally – act in line with this consensus. Liberal neutrality, in this case, would not be violated.

## 2. Anti-Liberal Criticism

Why, then, is liberalism still considered a blight in some feminist circles? In the following sections, I explicate three subsets of feminist criticisms of liberalism which fall within a broader critique. These arguments are, briefly: 1. Liberalism's focus on selves as individuals is ontologically flawed – call this the *Ontological Objection*; 2. Liberalism's focus on selves as individuals makes it so that the theory can't detect harms like oppression that arguably are intrinsically targeted as groups (rather than individuals) – call this the *Epistemic Objection*; and 3. Liberalism's focus on selves as individuals erases the differences between persons – call this the *Identity Objection*. I will first (briefly) explain the broader critique in order to situate the three specific criticisms.

The broad critique, the *Individualism Critique*, maintains that the fundamental problem with liberalism is a misguided focus on *individuals*: individuals who decide to enter social relationships only as a matter of last resort; isolated, self-sufficient individuals who, motivated by fear, grudgingly decide to come together to protect themselves and their interests. Seyla Benhabib and others draw our attention to Hobbes, who (in)famously argues that men are like mushrooms, springing up from the earth as fully grown and independent individuals (Benhabib 1992; cf. Schwartzman 2006: 166).<sup>5</sup> This sort of individualism is seen as both fictional and inimical to the lives of women and men, who are always situated within societies, communities and families.

---

<sup>5</sup> It's worth noting that Hobbes is not a liberal theorist, but a social contractarian. Rousseau, too, fits the social contractarian mold – and he is also not a liberal. The simple equation of liberalism and social contractarianism is a weak point for some anti-liberal feminist critiques, but this is a topic for another paper.



## ***2.1. The Ontological Objection***

The best-known version of this critique is ontological; feminist critics (like Iris Marion Young and Alison Jaggar) join with communitarians (like Michael Sandel and Charles Taylor) to charge liberalism with relying on a *false* conception of the self. I take the *Ontological Objection* to run something like this: Liberalism views selves as individuals, and thus does not view them as members of families, groups, or societies. But selves are not isolated individuals; they are always members of families, groups, and societies. Liberalism thus has a flawed conception of the self, and therefore we should reject liberalism as a viable theory.<sup>6</sup> The *Ontological Objection* is a fairly straightforward metaphysical reading of the *Individualism Critique*, and can be summed up as follows:

**Ontological Objection: Liberalism rests on the mistaken view that selves are metaphysical individuals.**

In other words, the *Ontological Objection* takes issues with metaphysical individualism, which is perhaps best captured by citing certain feminist objections to it – objections which claim that people are (to some extent) constituted by their relationships. Jennifer Nedelsky, for example, argues that “it is the very nature of human selves to be in interaction with others. In important ways, they do not exist apart from these relations” (Nedelsky 2011: 55).

We should immediately contrast metaphysical individualism with normative individualism, which is captured by liberalism’s focus on the individual as the primary unit of analysis. In other words, liberalism is, indeed, committed to the analytic priority of individuals

---

<sup>6</sup> Carol Hay explains this objection nicely - see Hay 2013: 23.

(Nussbaum 1999). But this commitment says nothing about the nature of individuals as such. Instead, it only requires that individuals come first when we are considering the relationship between the state and its citizens; no metaphysical claim is involved. I will return to this topic later when I address the *Ontological Objection* in more detail.

The crux of the *Ontological Objection* is the first premise (that a focus on selves as individuals precludes a focus on social selves), but different critics spell out this premise differently. Some early critics, like Iris Marion Young and Alison Jaggar, have criticized liberalism for thinking of individuals as ontologically prior to social groups (Jaggar 1983: 28; Young 1990: 44). Young argues that this (perhaps implicit) social ontology then gives way to a flawed normative conception of selves as individuals (Young 1990: 45), and Jaggar refers to liberalism as political solipsism (Jaggar 1983: 40).<sup>7</sup> Lynne Arnault has agreed with Young and Jaggar, and is worth noting here for her artful description of liberal selves as “rational, solitary, interest-driven monads whose every interpersonal action is an effort to maximize personal interests” (Arnault 1989: 192). The reality, for these feminists, is that selves are not isolated individuals. Positing that selves are individuals amounts to denying the way in which people *actually are* constituted by their traits and relationships.

It is important to note, however, that Rawls’ “Justice as Fairness: Political not Metaphysical” was not published until 1985, and *Political Liberalism* was not published until 1993. Since I am explicitly defending a political liberal position, it seems unfair to criticize feminists for taking Rawls’ earlier liberalism to task. Moreover, so much good work has already been done in response to the early version of this critique. Martha Nussbaum, Ruth Abbey, Lori

---

<sup>7</sup> See also Hirschmann 1994: 1861-8; Smiley 2004: 1602-6

Watson and others have already defended a political liberal – feminist – position against this critique, and they have done so quite well.<sup>8</sup>

While some political liberal feminists may think that Nussbaum’s work decisively ends this debate, the same charges leveled by Jaggar and Young *continue* to be leveled by anti-liberal feminists to this day. Linda Alcoff, for example, has recently suggested that liberalism – and specifically the individualism at its center – is (at least partially) to blame for the difficulty women have when attempting to prosecute their husbands or dates for rape. She argues, further, that contractarian approaches (of which, on her account, liberalism is one example) “may even provide support for victim-blaming” (Alcoff 2009: 129). She goes on to associate the (supposedly) problematic concept of consent directly with liberalism: “This invites us to wonder whether the use we make of the concept of consent and the difficulty we have with the concept of victim is peculiar to these liberal traditions. *Do these concepts harbor individualist metaphysical conceptions of selfhood and agency?* If so, what would formulations of sexual violence look like without *individualist conceptions of the self?*” (Alcoff: 2009: 126, emphasis added).

More recently, Nancy Hirschmann has argued that liberalism’s individualist metaphysics makes the theory both masculinist and ableist (Hirschmann 2013: 101). Rawls, she argues, betrays his metaphysical commitment to individualism when he talks about relationships as if they were freely chosen. He ignores the fact that “it is women’s attention to the realm of care that allows men to develop and believe in a conception of justice that operates on a foundation of rules, principles, rationality, and, of course, the logic of individual free choice” (Hirschmann 2013: 102). She continues her argument in the now familiar language of earlier feminist

---

<sup>8</sup> In addition to the collections in n.1, see: Nussbaum 1999; Hartley and Watson 2010; Watson 2010-2011.

critiques: “Rawls adopts a classically liberal abstract individualism as the foundation of his theory of freedom: as in social contract theory, the isolated individual is the center of the universe and the basic foundation; that individual decides which relationships and connections to forge and thereby creates his or her relationships and communities, which are thereby contractually produced on the basis of self-interest” (Hirschmann 2013: 104-5).<sup>9</sup>

Why do these charges continue to stick? Why does the *Ontological Objection* continue to appear in anti-liberal feminist attacks? Why has the liberal response, the straightforward denial of a metaphysical conception of the self, proved insufficient? To see why, we must further unpack the objection. At the heart of the *Ontological Objection* is the claim that there is something *dishonest* about liberalism. Liberalism’s supposed political (not moral) nature and its supposed neutrality are masks disguising the truth: the liberal state relies on a robust moral theory that champions an individualist conception of the self as a social atom. In her critiques of liberalism, for example, Catharine MacKinnon shifts back and forth between identifying her opponent as “liberalism” and “liberal morality”; she attacks what she calls “liberal meanings of personhood” (MacKinnon 1989: 209). On MacKinnon’s view, even though liberalism professes not to rely on any single moral theory (or comprehensive doctrine) in reality it most certainly does – and this single moral theory is inimical to women. So while liberals may say that their theory does not rely on any specific conception of persons, they are either lying or mistaken; liberalism most certainly does rely on a metaphysically robust conception of selves – and this conception is fundamentally flawed.

Like Hirschmann, some feminist critics continue to argue that liberalism’s conception of the person is implicitly, even fundamentally, masculine. Kathy Miriam, for example, takes

---

<sup>9</sup> It is worth noting that Hirschmann’s references to Rawls are almost exclusively to *Theory of Justice*, before his political liberal turn.

liberalism to task for its (again, supposed) individualist conception of the self. She supports Pateman's argument in *The Sexual Contract* that the domination of women is actually built into the structure of liberalism itself, and thus liberalism cannot exist without oppressing women (Pateman 1988: 6). Miriam argues that there is an "inextricable relationship between female *agency* on the one hand, and male *domination*, on the other, in a liberal social order [...] Male mastery and female subjection is a power relation structured into liberalism (Miriam 2005: 10). Like Alcoff, she links the problem with liberalism to its focus on the individual (Miriam 2005: 2).

What this discussion highlights is the fact that the *Ontological Objection* is not simply that liberals have a flawed ontological conception of selves, but rather that this flawed conception has serious consequences. Central to the objection is the phenomenological thought, also emphasized by communitarians like Sandel, that we cannot recognize ourselves in liberalism's picture of us (Sandel 1982: 179-83). In order to understand ourselves, or recognize ourselves, we have to imagine ourselves with certain constitutive commitments. Ronald Dworkin explains the charge, that "people need community not only for culture and language but for identity and self-reference, because people can only identify themselves, to themselves, as members of some community to which they belong" (Dworkin 1989: 488). When liberals claim not to have a metaphysical conception of the self, they turn political theory into something concerning unrecognizable creatures called individual 'citizens,' not something about socially situated people like us. This may, in turn, have the further consequence of turning liberal neutrality into a mask disguising the truth – that liberal individualism is a mask for something decidedly non-neutral, and deeply morally flawed.

## 2.2. *The Epistemic Objection*

In this section, I turn to look at a specific subset of objections which criticize an implication of liberalism's normative individualism: those which claim that liberalism lacks adequate means of detecting group harms. Because these objections are about what liberalism cannot adequately detect, I group them together as the *Epistemic Objection*: that liberalism's focus on individuals makes it so that the theory cannot detect group harms like oppression. The general structure runs something like this: If a theory focuses on the individual instead of on social groups, it can't take account of structures of power that affect people based on group membership. Group harms like oppression accrue to people based on group membership, and thus cannot be detected through an individualistic framework (Schwartzman 2006, 2013).<sup>10</sup> A brief restatement of the *Epistemic Objection* runs as follows:

**Epistemic Objection: Liberalism cannot adequately detect group harms like oppression.**

If this charge sticks, it deals a heavy (perhaps even fatal) blow to liberalism. Surely a political theory should be able to address oppression - especially liberalism, which is self-avowedly concerned with the freedom and equality of all citizens.

Lisa Schwartzman explains the problem in the following way: "Because liberal theory grants rights to individuals *as individuals*, it primarily recognizes violations of rights that occur one at a time, to individuals *as individuals*" (Schwartzman 2006: 27; cf. Schwartzman 2013: 46).<sup>11</sup> Oppression, however, does not impact people *as individuals*; it only impacts people

---

<sup>10</sup> Carol Hay has an excellent discussion of objections of this sort. See Hay 2013: 24.

<sup>11</sup> Schwartzman concedes that liberalism can eventually take account of oppression by pointing out that oppressed individuals have different opportunities, for example, than non-oppressed

insofar as they are members of a specific group (the group Women, for example) (see e.g. Cudd 2006). Iris Marion Young makes the same point when she argues that individualist theories cannot acknowledge group harms: "The discourse of liberal individualism denies the reality of groups [...] Without conceptualizing women as a group in some sense, it is not possible to conceptualize oppression as a systematic, structured, institutional process" (Young 1997: 17).<sup>12</sup> Other non-liberal feminists who have leveled this charge include Catharine MacKinnon (1989), and Carole Pateman (1988).

There are, then, two related aspects of the *Epistemic Objection*. There is the primary, epistemic dimension - because of its individualism, liberalism does not have adequate means of detecting group harms like oppression. But there is also a practical dimension - because it cannot properly *detect* oppression, liberalism cannot adequately *address* it. If this is really the case, then so much the worse for liberalism.

### **2.3. *The Identity Objection***

The final objection that I will consider is the *Identity Objection*, which charges that - because liberalism problematically rests on a notion of all people as free and equal, and because all people are granted the same bundle of rights - liberalism must assume that all people are the same. Anti-liberal feminists have raised two related problems with this assumption. First, there is a descriptive problem - it is simply not the case that all people are equal. When liberals claim that all people *are* in fact equal, they are obscuring the reality that many people are, in fact, *not*. I think this claim can be readily dismissed; it is a straightforward case of confusing description

---

individuals. But even though liberalism might ultimately acknowledge oppression, it is not a sufficient theory for discovering instances of oppression; radical critiques are necessary for this task.

<sup>12</sup> For Ruth Abbey's discussion of this point, see Abbey 2011: 16.

with prescription. Liberalism does not assume that people are all, descriptively, equal. Liberalism requires that people (as citizens) are treated as moral equals; it is a normative ideal.

I focus instead on the second, normative problem - that liberalism, with its universalistic conception of individual rights, requires that different identities be obscured; when liberalism assumes that people are equals, then, it is really assuming that they are all *the same*. There is both a general and a more specific way to spell out this problem. Some feminists charge that liberalism's focus on equality requires the elimination of difference, broadly construed (Alcoff 2006).<sup>13</sup> Linda Alcoff, for example, argues that "visible difference threatens the liberal universalistic concepts of justice based on sameness by invoking the specter of difference" (Alcoff 2006: 180). The more specific problem, applying to relationships between men and women in particular, is that liberal equality requires women to become more like men in order to receive the formal equal treatment that liberalism professes to afford them (Schwartzman 2006: 100). A brief restatement of the *Identity Objection* runs as follows:

**Identity Objection: Liberalism requires a political definition of persons that obscures differences between them.**

But the criticism here goes deeper; the problem isn't just that liberalism erases real differences between persons. The *Identity Objection* can also be understood as cutting to the

---

<sup>13</sup> See also Young 1990: 229; Zerilli 2009: 298; Mills: 2011: 584. Young argues: "Liberal individualism denies difference by positing the self as a solid, self-sufficient unity, not defined by anything or anyone other than itself. Its formalistic ethic of rights also denies difference by bringing all such separated individuals under a common measure of rights." The idea here is that granting the same set of formal rights to all people problematically blurs the distinctions between them. If liberalism were to recognize difference, it would no longer be able to offer equal rights to all; instead, it would have to offer different rights to different groups of people, as necessary.



very heart of liberalism. Recall that one of the two traits of liberalism (discussed above) was its *neutrality*; liberalism refuses to endorse one comprehensive doctrine over another. The *Identity Objection* makes it look like this neutrality is only nominal. Why? It's helpful here to refer back to Rawls. In Rawls's hypothetical contract scenario, the original position, there are certain features of persons that must be bracketed. In order to ensure equal consideration for all comprehensive doctrines, any personal attributes that are morally arbitrary should not factor into our decision regarding the principles of justice. People in the original position are behind what Rawls calls the veil of ignorance - it is this feature that ensures we do not take morally arbitrary biases into account when determining principles of justice for our society.<sup>14</sup> Rawls argues: "The parties [in the original position] are symmetrically situated with respect to one another and they are in that sense equal; and what I have called 'the veil of ignorance' means that the parties do not know the social position, or the conception of the good [...] and much else, of the persons they represent" (Rawls 2005: 305).

Anti-liberal feminists have highlighted two problems with this approach. First, to what extent are these features really arbitrary? And second, how do we decide which features to bracket? According to some radical feminist critiques, it is the people in positions of power who determine what features are arbitrary; the supposedly "symmetrically situated" persons in the original position are really just representations of the powerful majority. Thus, the perspectives of the powerful come to be seen as what is neutral. Catharine MacKinnon and Lisa Schwartzman endorse this reading of liberalism. Schwartzman explains: "Because the liberal approach to sexual equality requires that people *already be the same* before they can be seen as

---

<sup>14</sup> Note that calling parties in the original position "people" is a bit misleading. In *Political Liberalism*, Rawls claims that parties in the original position are "rationally autonomous representatives of citizens in society" (Rawls 2005: 305). His wording here highlights, again, liberalism's focus on citizens instead of persons as such.

deserving of equal treatment, women who suffer the effects of sex discrimination encounter difficulty achieving equality" (Schwartzman 2006: 100). The bracketing of certain traits of persons as a method to achieve equality is, in essence, the demand that we view all people as exactly the same before we can treat them equally.

For Schwartzman and others, it is the drive to eliminate arbitrary features of persons that leads to this problematic result. Liberalism's goal, to treat all people as equals, ends up blurring very real differences between persons, dubbing these differences 'arbitrary.' Carol Hay explains the problem in the following way: "Ignoring the differences between people perpetuates oppression because it allows the ideals of the majority to be presented as neutral and universal. Liberal abstraction becomes functionally assimilationist, presenting the ideals and concerns of those in power as if they were the ideals and concerns of everyone" (Hay 2013: 20). In short, liberal equality requires blurring the distinctions between persons; it requires that we see all people as the *same* if we are to understand them as *equals*.

### **3. A Political Liberal Response**

In this section, I turn to a defense of political liberal feminism from the three types of objections I've just discussed. I will be stressing a key point throughout my response: liberalism is not a full-fledged moral or metaphysical theory. *Liberalism is fundamentally about the relationship between a state and its citizens*, and political liberalism is founded on the very idea that it shouldn't be an exhaustive moral theory. To criticize liberalism for being a flawed moral theory, or to criticize it for resting on flawed metaphysical concepts, is to criticize something that needn't be attributed to liberalism at all.

### *3.1. Response to the Ontological Objection*

Recall the brief formulation of this objection:

**Ontological Objection: Liberalism rests on the mistaken view that selves are metaphysical individuals.**

I take it that what critics are objecting to here is the analytic priority of individuals, which is, indeed, a part of liberal theory. But note that the analytic priority of individuals only requires that individuals come first as a unit of analysis when we are considering what sort of relationship the state should have to its citizens. Political liberals argue that to say this is not to make a metaphysical claim about the nature of persons. It is worth quoting Rawls at some length here: “In this case the conception of the person is [...] adapted to a political conception of justice and not to a comprehensive moral doctrine. It is in effect a political conception of the person, and [...] a conception of citizens. Thus, a conception of the person is to be distinguished from an account of human nature” (Rawls 1999: 397n.15). Will Kymlicka expands on Rawls’ point when he explains that, for Rawls, the political conception of the person isn’t “intended as a general account of the relationship between the self and its ends applicable to all areas of life, or as an accurate portrayal of our deepest self-understandings” (Kymlicka 2002: 235).<sup>15</sup> In other words, when we are talking about liberalism, we are talking about a political theory that considers persons qua citizens, and not persons as such (Kymlicka 2002: 235; Talisse 2012: 132). It seems no ontological view of individuals is required.

---

<sup>15</sup> Kymlicka also notes that the “veil of ignorance is not an expression of a theory of personal identity” (2002: 63).

This may all sound like old hat to political liberals, but note that – for critics – there is an immediate rejoinder. Critics can grant that liberalism is a theory that focuses on the relationship between the state and its citizens. But they can then argue that *even saying this much* is proposing a metaphysical conception of the self. Liberalism, by focusing on the relationship between the state and its citizens, presupposes a special office – citizen – that selves hold, which is distinct from selves as such. That is, political liberalism presupposes that the role of citizen is detachable from any broader theory about what we are. In other words, the fact that liberals think it makes sense to distinguish ‘persons qua citizens’ from ‘persons as such’ demonstrates *already* that they make a claim about ontology – that selves are distinct from the roles they play. Indeed, this point seems to be made plain in Rawls’ representational device, the original position (see Mulhall and Swift 2003: 464-5).

Moreover, Rawls himself explicitly builds certain commitments into his political conception of persons as citizens. Citizens, for Rawls, are free in three ways: first, they conceive of themselves and their fellow citizens as capable of forming, revising, and rationally pursuing their own conception of the good. Second, citizens view themselves as “self-authenticating sources of valid claims” (Rawls 2005: 32); in other words, they see themselves as entitled to make claims on other citizens and institutions so that they might advance their own conception of the good. Finally, citizens can take responsibility for their ends; they can adjust their goals to what they are entitled to receive (Rawls 2005: 30-34). These three traits help to specify what Rawls means when he says that citizens are free and equal in their possession of a capacity for a sense of justice and a capacity for a conception of the good (Rawls 2005: 19).

Merely claiming that none of this amounts to a metaphysical conception of personhood will not assuage the worries of feminist critics. So what can the political liberal say in response?

The first thing to note in response is that normative individualism is not based on the thought that individuals are all that matter, full stop. Normative individualism, instead, maintains that individuals are *ultimately* what matter, politically speaking - and this is why they must come first as units of analysis when determining the proper relationship between states and citizens (Anderson 2009: 132). What is the difference between the claim that individuals are the *only things* that matter, and the claim that individuals are what *ultimately* matter, politically speaking? There are at least two differences. The first difference is a matter of scope - normative individualism, as a tenet of liberalism, only applies within the political scope of liberalism itself. The second difference is that, to say that individuals are what ultimately matter is not to say anything about what, in turn, matters to individuals. Things like family and group membership can be hugely important to people - even partly constitutive of people - and there is nothing about normative individualism that denies this fact.

In addition, because of its commitment to neutrality, liberalism remains impartial regarding different metaphysical views about persons. There is nothing about normative individualism that precludes the possibility that selves are relational, for instance. Samantha Brennan helpfully points out: "It is true that rights attach to persons but there are a variety of different ways one might understand the nature of persons...If rights apply to persons, and the correct understanding of persons is as relational entities, then rights too are best understood as attaching to persons as relational entities" (Brennan 2004: 92). Normative individualism is non-committal regarding different metaphysical conceptions of the self (adopted by citizens), including metaphysical conceptions that feminists are likely to endorse.

As political liberals, we might be tempted, following Rawls, to claim that our theory proposes no conception of selves *at all*, but this is a bit disingenuous; in fact, it is likely to

bolster the claims of critics like MacKinnon, Miriam, and Alcoff who see something dishonest in liberal neutrality. Continuing to insist that, because liberalism remains quietist about the metaphysics of the self, it is not vulnerable to the *Ontological Objection* is merely talking past many feminist critics. Furthermore, continuing to highlight that the liberal conception of the self is *political* only plays into the critics' hands, since they argue that such a political conception only makes sense if a particular metaphysical conception of the self is true – regardless of what liberals claim.

The best way to respond to these worries is to draw upon our previous discussion of neutrality and distinguish between *rejecting* all conceptions of personhood and *declining to endorse* any such conception. Importantly, liberalism does not reject all conceptions of personhood; it does not rest upon the faulty assumption that we need not know what a person is in order to do political theory. It does not reject conceptions of personhood that focus on the social aspects of identity. But liberalism does decline to endorse any *particular* conception of the person. Moreover, this lack of endorsement does not force liberals into a default individualist position; by declining to endorse, say, constitutively relational conceptions of personhood, the liberal is not thereby arguing for the contrary individualist conception. Perhaps, then, instead of understanding liberalism as a theory resting on no metaphysical conception of the self at all, we should understand it as a theory that could potentially rest on any metaphysical conception of the self.

Indeed, this way of viewing things is supported by Rawls' idea of an overlapping consensus, which he defines as consisting of “all the reasonable opposing religious, philosophical, and moral doctrines likely to persist over generations and to gain a sizable body of adherents in a more or less just constitutional regime” (Rawls 2005: 15). Citizens are meant to

endorse a certain theory of justice from whatever comprehensive doctrine they happen to hold (see e.g. Rawls 2005: 150-1). This same idea can apply to conceptions of selfhood; liberal theory allows citizens to support theories of justice which rest on any conception of selfhood they happen to hold (as long as this conception is reasonable).<sup>16</sup> So while liberalism does decline to endorse social conceptions of selfhood, it similarly declines to endorse individualist conceptions.

For this reason, the further phenomenological charge that we are unable to see ourselves in the liberal picture of citizenship is unfounded. Much of the phenomenological worry arises from Rawls' original position, where individual (representatives of) citizens reflect upon principles of justice behind a veil of ignorance. How could we possibly imagine ourselves in this scenario? Moreover, doesn't the *very idea* that citizens could deliberate individually in this way presuppose an individualist conception of the self? Rawls himself recognizes this worry. "The description of the parties [in the original position] may seem to presuppose a particular metaphysical conception of the person," he notes (Rawls 2005: 27). But the parties in the original position are not citizens – they are representatives of citizens in a hypothetical scenario. We are meant to be able to imagine ourselves in this position, but this doesn't mean that at any given moment we should be able to picture ourselves as disembodied, featureless beings. Perhaps we can put the point colloquially: imagining ourselves in the original position is much like taking into account the opinions and beliefs of our fellow citizens. How would we think about this principle of justice if we were differently situated? Could we justify it? Note that,

---

<sup>16</sup> It's important to remember here that reasonableness is not a substantive concept. For Rawls, reasonable persons are "ready to propose principles and standards as fair terms of cooperation and to abide by them willingly, given the assurance that others will likewise do so" (Rawls 2005: 49). Unreasonable people are unwilling to honor or propose "any general principles or standards for specifying fair terms of cooperation. They are ready to violate such terms as suits their interests when circumstances allow" (Rawls 2005: 50).

again, no metaphysical conception – including an individualist conception – of the person is required.

### ***3.2. Response to the Epistemic Objection***

Recall the brief formulation of this objection:

**Epistemic Objection: Liberalism cannot adequately detect group harms like oppression.**

As we have seen, nothing about normative individualism precludes an understanding of persons as relational or as members of certain groups, thus, there is nothing about it that prevents liberals from detecting oppression. In fact, many liberal feminists point out that normative individualism is actually helpful for understanding and addressing oppression. Ruth Abbey and Carol Hay, for example, both cite liberalism's commitment to viewing all citizens as free and equal persons, arguing that any society which systematically oppresses a group of people fails to meet the standard for equal respect (Abbey 2011: 217; Hay 2013: 21). Abbey then goes on to argue that normative individualism does not prevent liberal feminists from viewing individual women as members of the group Women, nor does it prevent liberal feminists from taking account of differences among women. "Contemporary feminist liberals," she argues, "do not assign women to a single oppressed category by virtue of their biological features nor any essence of womanhood [...] Their category of women is political" (Abbey 2011: 267).

But my response, so far, hasn't gone far enough to assuage non-liberal feminists. Here's why: the anti-liberal feminists can't just be claiming that political liberals have no concept of oppression (and that, therefore, they can't address it). There is no doubt that political liberals



isolate certain phenomena and call them ‘oppression’. What the critics must be arguing is that, regardless of what liberals *say*, *real* cases of oppression remain invisible to them; in other words, liberals can’t recognize certain kinds of oppression *as oppression*. Liberalism’s commitment to individualism renders liberals incapable of detecting certain kinds of oppression, and it is exactly these kinds of oppression that they, therefore, can’t address. In fact, this highlights what is particularly insidious about liberal “neutrality” – one might think isn’t neutral at all, but in fact obscures the structures of male domination that serve to continually oppress women (see e.g. MacKinnon 1989).<sup>17</sup>

The best way that I can see to address this objection is to offer a series of counterexamples in which political liberals both converge with anti-liberal feminists on specific cases of oppression and recognize them *as oppression* using the tools that liberalism provides. First, I will isolate cases in which anti-liberal feminists conclude that there is oppression, particularly a type of oppression that liberals can’t detect. I will then demonstrate that, in actuality, liberals *can* recognize most of these cases as cases of oppression.

Lisa Schwartzman is one of the most vocal recent representatives of the *Epistemic Objection*. She charges that, because of liberalism’s methodological individualism and its neutrality, the theory is unable to detect *clear cases* of oppression like sexism, racism, classism, and heterosexism (Schwartzman 2006: 165). Specifically, Schwartzman considers pornography as a practice which instantiates several different forms of social inequality (including the four just mentioned). The only way to understand pornography as a site of oppression is to begin from the perspective of women’s subordination and ask whether or not social practices like pornography “reinforce, perpetuate, or contribute to women’s inequality” (Schwartzman 2006:

---

<sup>17</sup> For a good discussion, see Brake 2004.

165).<sup>18</sup> Liberals, according to Schwartzman, with their focus on individuals and neutrality, can only look at whether a specific individual has consented; they cannot look at the effects that pornography on society as a whole, and on subordinated social groups like women in particular. Thus, for Schwartzman, even though liberals could potentially recognize pornography as a harm, they cannot recognize it *as oppression*, due to their focus on individuals and neutrality.

The first thing to note, again, is that neutrality enters the game at the justificatory stage, so it's not clear that Schwartzman's charge applies to liberalism's neutrality. But the charge might still stick to liberalism's (supposed) individualism – even if we grant my earlier point that this individualism is normative. But, as I've shown, normative individualism does not preclude an understanding of citizens as social beings; the sort of individualism at the center of liberalism does not, for example, prevent seeing a particular woman as part of the group Women. In the case of pornography, if it is the case that the practice subordinates women, then it arguably fails liberalism's own requirement to see all citizens as free and equal (Watson 2007: 476).

Can liberals accept that pornography subordinates women? Certainly. Pornography might be seen as a type of propaganda supporting the degradation of women as sex-objects. Joel Feinberg, for example, notes that pornography “reinforce[s] macho ideology” and does “manifest harm” to women (Feinberg 1985: 153). Liberals can understand this manifest harm as a straightforward denial of equal treatment – treatment that is demanded by liberalism itself (Langton 1990: 353). Liberals are not forced, contra Schwartzman, to focus narrowly on questions of individual consent (Schwartzman 2006: 164). Insofar as pornography denies equal treatment to women (which we can stipulate for our purposes here) it denies equal treatment to

---

<sup>18</sup> Schwartzman thus continues the ‘civil rights approach’ of Catharine MacKinnon and Andrea Dworkin, who define pornography as “the graphic sexually explicit subordination of women” (MacKinnon and Dworkin 1997: 435). See Watson 2007, 2010.

individual women qua woman. If we grant that pornography subordinates women, then liberalism can certainly prohibit the practice on the grounds that it denies key tenets of liberalism.

I will consider one more example. Will Kymlicka's work on multiculturalism is explicitly concerned with the oppression of national minority groups like American Indians and Canadian First Nations. Note that he addresses these cases as cases of oppression – not just moral wrongness or injustice (although these might also be the case). He also clearly understands this oppression in a liberal framework, which emphasizes loss of autonomy for individual members and historic injustices against these groups. What makes his account distinctively liberal is that membership in these cultural groups cannot be mandated by the state, but instead must be a matter of self-identity. Individual members must also have exit rights available to allow transition from any identity group, and the groups must not violate the basic civic and political rights of their members (Kymlicka 1998: 147). In fact, it is this latter part that illustrates the particular strengths of a liberal approach. Kymlicka's theory provides a great example of the liberal focus on individual rights, and demonstrates that this focus is not incompatible with group-specific rights or policies.<sup>19</sup>

There is a final point to mention in response to the *Epistemic Objection* which I only have space to mention, not defend in full. This point, defended by some liberal feminists, is that liberalism actually has *better* ammunition to attack oppression than do other political theories. The most often-cited reason of why this might be the case is the reason I have already mentioned - that liberalism demands equal respect for all citizens, no matter what conception of the good

---

<sup>19</sup> See also Kymlicka 1995, 2009.

they choose to endorse. Liberalism has at its disposal powerful resources for detecting and addressing oppression, in spite of claims to the contrary. Whether or not it is the best vehicle for ending oppression remains here an open question.

### ***3.3. Response to the Identity Objection***

Recall the brief formulation of this objection:

**Identity Objection: Liberalism requires a political definition of persons that obscures differences between them.**

Earlier, I discussed two points related to this objection that I wish to respond to in this section: first, that liberalism assumes that all people are equal when they are clearly not; and second, that liberalism either obscures the differences between persons, broadly speaking, or between men and women more specifically. The first criticism is easily overcome by pointing out that liberalism does not, in fact, assume that all people are equal (descriptively speaking). Rawls explicitly draws the distinction between what he calls the moral agent as a "free and equal citizen as a member of society" and a moral agent "in general" (Rawls 2005: 109). There is nothing which mandates that liberalism view all people as descriptive equals. Equality is a normative requirement for people understood as liberal citizens. Insofar as people are not descriptively equals - something that liberalism can certainly recognize (see my response to the *Epistemic Objection* above) - the liberal state is failing to meet its own ideals.

The second point, I think, confuses an ideal of *equality* with an ideal of *sameness*. I've appealed throughout this discussion to Rawls' insistence that we bracket morally arbitrary traits;

this feature of Rawls is often cited as either an example of a liberal promotion of sameness, or an example of liberalism's tendency to mask injustices. Lisa Schwartzman, for example, notes that Rawls "describes the 'initial social position' into which one is born as 'arbitrary'" (Schwartzman 2013: 53). She cites Rawls, who in a *Theory of Justice* argues that "the natural distribution [of talents] is neither just nor unjust; nor is it unjust that men are born into society at some particular position. These are simply natural facts" (Rawls 1971: 102, cited in Schwartzman 2013: 53). She goes on to argue that social circumstance is not, in fact, random, arbitrary, or natural. "The social and economic forces that create and maintain structures of power and oppression – including class, race, and gender – are not governed by luck or accident" (Schwartzman 2013: 53).

But again, what liberalism requires is that we eliminate the influence of *morally* arbitrary traits on justice considerations *from the perspective of the state*. So, from the perspective of the state, we *are* all the same in one respect: we all deserve respect as free and equal citizens. But this sameness does not bear on any substantive facts about us as persons. Admittedly, at first blush it might seem entirely wrongheaded of Rawls to argue that there is nothing just or unjust about being born into one particular social position over another; surely he can only be committed to this position because he thinks personal traits, like race and sex, are arbitrary. But the charitable way to interpret Rawls' claim is to understand that *there is* nothing just or unjust about the *fact itself* that we are born into a given social position instead of another. This is just a natural fact. What makes it just or unjust is that social institutions treat us differently based on natural facts which should have absolutely no bearing on our status as free and equal citizens – this is just what it means for a trait to be morally arbitrary from the perspective of the state.

It is worth noting here that if we look at comprehensive or perfectionist liberal theories, we might be tempted to say that liberalism does adopt certain substantive views about people *qua people* (depending on how deep the comprehensive-ness goes). Indeed, this is why some comprehensive liberal accounts make such good targets for non-liberal feminists. This chapter defends only political liberalism. But I will mention that, *insofar as comprehensive liberals are liberals*, they are still, at bottom, dealing with the relationship between the state and its citizens.

#### **4. Conclusion**

In this chapter I have illustrated part of what I take to be a methodological mistake at the center of debates between political liberals and non-liberal feminists. Non-liberal feminists often object to liberalism on the grounds that it relies on a problematic conception of selves as isolated individuals; these critiques make up the broader critique that I've referred to as the *Individualism Critique*. I've focused on three versions of this critique, the *Ontological Objection*, the *Epistemic Objection*, and the *Identity Objection*, and I've offered initial responses on behalf of the political liberal. The heart of my response is that, contrary to what critics argue, political liberalism does not rely on a problematic conception of the self because its conception of the self is political, that is, lacking any metaphysical depth. Liberalism is a political theory which focuses on the relationship between a state and its citizens, and as such relies on an understanding of people as citizens and not of people as such.

I want to close by highlighting what I take to be at stake in this debate. Political liberalism is not metaphysically individualistic, as I've mentioned here, and thus doesn't fall prey to the *Individualism Critique*, as I've argued. But it does rely on normative individualism. Political liberalism does require that individuals come first as a unit of analysis, and it does

understand harms like oppression as harms that accrue to groups *made up of individual persons*. And it is true that liberalism applies important normative requirements (like rights) to individuals, as well. Even being individualistic in these ways will likely feed the opponents' fire, and I recognize that. But the political liberal position that I have defended here has a particular strength in its normative individualism. Political liberalism demands that we not lose the individual woman against the backdrop of oppressive social structures. It is true that oppression accrues to me, as a woman, based on my membership in the group Women. But it accrues *to me*, as well as to us.

The liberal insight that I am to be treated as a free and equal citizen, no matter what sex, race, sexuality, or class I am is a powerful tool in combatting oppression. The moral equality of all people which forms the very foundation of liberalism points out oppression as a serious injustice and demands that it come to an end. The liberal thought that the state should not be granted the power to impose normative standards on one's life is similarly a powerful feminist tool. I am free, as a liberal citizen, to form my own self-interpretation and my own conception of the good life. Of course, this includes the idea that freedoms can (and should) be limited to prevent harm to others. It is perfectly legitimate for the state to curb one's behavior to prevent harm to others (and yes, this requires careful specification of what counts as harm to others that is sufficient for the state to be entitled to prevent it). This is a hugely important topic, but one for another paper. What I hope to have illustrated here is that, according to political liberalism, I am free to live according to my personal values just as all women, and all liberal citizens, are likewise free. It is in this way, I think, that political liberalism is a powerful feminist ally.

## CHAPTER II

### TRADITIONAL AUTONOMY: HIERARCHICAL, HISTORICAL, AND NORMATIVE COMPETENCE THEORIES

#### **1. Distinctions and Classifications**

There are, broadly speaking, three different axes that help us analyze theories of autonomy. We can place a theory on the Procedural/Substantive spectrum, we can place it on the Internalist/Externalist spectrum, and we can place it on the Local/Global spectrum. A theory is *procedural* if it defines autonomy in terms of the process an agent undergoes when considering whether or not to undertake a specific action, or whether or not to have a specific desire or belief. A theory is *substantive*, on the other hand, if it defines autonomy in terms of the content of the agent's considered action, desire, or belief. Procedural theories, which do not consider this content, are thus often called content-neutral. Between these two extremes - of purely procedural theories on the one hand and purely substantive theories on the other - are theories which contain a combination of procedural and substantive requirements for an agent's autonomy. Such theories might be weakly substantive, for example.

Autonomy theories can also be placed along the Internalist/Externalist spectrum. A theory is *internalist* when it appeals to an agent's internal psychological processes or states to determine whether or not the agent is autonomous. A theory is *externalist* when it appeals to factors in the external world - factors outside of the agent - pertaining to the situation in which an agent finds herself. In other words, those who endorse internalist accounts think that autonomy is a psychological feature of agents. To use Diana Meyers' helpful language, autonomy is something a person *has* (Meyers 1987). In order to be autonomous, a person's psychology must be in the right order, or must meet certain standards. On externalist accounts,



however, a person's psychology can be in the right “order” while she nonetheless fails to be autonomous. Certain external features have to obtain in order for an agent to achieve autonomy. Once again using Meyers' language, if we endorse an externalist theory of autonomy, we might say that autonomy is something that *happens to* an agent. Just as with the Procedural/Substantive spectrum, a theory can combine both internalist and externalist factors as relevant to determinations of autonomy.

Finally, we can place autonomy theories along the Local/Global spectrum. Most theories of autonomy that I will consider are local accounts, that is, they consider the autonomy of specific desires, preferences, actions, or choices. Global theories, on the other hand, consider the autonomy of an agent over a period of time. Marina Oshana is one of the leading proponents of a global theory. She describes a global theory as one that construes autonomy as “the property of a person having *de facto* power and authority over choices and actions significant to the direction of her life” (Oshana 2006: 2). Strictly speaking, local autonomy applies to choices while global autonomy applies to people. For most of my discussion, however, I will speak of autonomous people as those who make autonomous choices. If Agent A, for example, autonomously makes choice C, I will speak of A as being autonomous in this situation. In other words, unless otherwise specified, I assume a local account of autonomy.

## **2. Hierarchical Theories**

Most of the traditional theories that feminists criticize are procedural and internalist. But characterizing a theory as procedural and internalist is only a broad characterization. There are, in fact, many different theories that fit this description. In order to understand the feminist critiques, we need to have a better grasp on the traditional theories themselves. In the following

section, I will present representative views of different sorts of autonomy theories. I will discuss Harry Frankfurt's theory as a representative of hierarchical theories, John Christman's theory as a representative of historical theories, and Gary Watson's theory as a representative of normative competency theories. I present objections that have been raised against each of these accounts in the mainstream literature. In the next chapter, I will discuss feminist objections.

### ***2.1. Harry Frankfurt***

In his 1971 essay, "Freedom of the Will and the Concept of a Person," Frankfurt presents a hierarchical theory of autonomy according to which people are autonomous if they have a second-order volition that a first order desire be effective. Note that there are two requirements; first, an agent must have a volition that a particular first order desire of hers be effective in her decision making. If Agent A has an effective first-order desire to  $\phi$  - that is, she *wills*  $\phi$  - she must have a second-order desire to want to desire  $\phi$  - that is, she must have a second-order desire that she will  $\phi$ . But, second, Agent A must also want her first-order desire to move her to action; she must want her first-order desire to  $\phi$  to be effective (Taylor 2005: 4). Using Frankfurt's terminology, she must have a second-order *volition* that her first-order desire move her to  $\phi$ . The two conditions required for autonomy are thus the 'Desire condition' and the 'Volition condition'.

Consider Betsy and her strong desire to eat a pint of Ben and Jerry's Brownie Batter ice cream; she has a strong first-order desire to eat it. But Betsy also knows that there are 1200 calories in a pint of Brownie Batter ice cream, and that it's probably not healthy to eat 1200 calories of ice cream all at once. She attempts to tell herself that she doesn't *really* want the ice cream, but fails to convince herself. Her second-order desire to not want to want to eat the ice

cream fails to override her first-order craving. She eats the ice cream and feels horrible about it afterwards.

Betsy thus fails the Desire condition (she does not have a second-order desire that supports her first-order desire for ice cream) and the Volition condition (she does not have a volition that her first-order desire move her to act). She does not *want to want* the Brownie Batter, and so her decision to eat it is heteronomous. If, on the other hand, Betsy decides that she deserves a pint of ice cream after her long work day - her second-order desire might change. She may now want to want to eat the ice cream. According to Frankfurt, Betsy would now be autonomous regarding her decision.

## ***2.2. The Regress/Arbitrariness Problem and Wholeheartedness***

This result is fairly intuitive. Betsy's conflict in the first scenario, and her resulting regret, might signal to us that she has not chosen autonomously, while the ease with which she decides in the second scenario, and her happy satisfaction afterwards, might signal to us that she has.<sup>20</sup> But there are problems with Frankfurt's initial 1971 theory, the most serious of which is the *Regress/Arbitrariness Problem*.<sup>21</sup> Hierarchical theories are so named because they posit a hierarchy of desires - we have first-order desires and second-order desires regarding our first-order desires. But we are left wondering whether or not these second-order desires are autonomous. If it is the second-order volition that determines our autonomy-status, this second-order volition should also be autonomous.<sup>22</sup> But Frankfurt's theory provides no answer to this question. If we are meant to

---

<sup>20</sup> Note that signaling autonomy is not the same as constituting autonomy.

<sup>21</sup> For discussions of this objection, see Mackenzie and Stoljar 2000; Taylor 2005; Oshana 2006; Christman 2007.

<sup>22</sup> Michael Bratman argues, for example, that Frankfurt reduces an agent's identification with a desire to an agent's relevant higher-order attitude concerning that desire (2002: 66).

simply stop analyzing after the second-order, then why? Why not stop at the first? Frankfurt's theory is threatened by either an infinite regress or an arbitrary stopping point.

In 1987 Frankfurt amended his theory to include a response to the *Regress/Arbitrariness Problem*. In "Identification and Wholeheartedness", Frankfurt provides a necessary condition for a second-order desire to be autonomous - an agent must wholeheartedly identify with it. As human beings, Frankfurt argues, we are particularly interested in what moves us to act. If we were not concerned with our motivations at all, or if they were to arise from something external to ourselves, there would be no way in which we could say that the motivations were *ours*. In fact, since our desires and motivations come to be "specially constitutive" of us as human beings, any action that results from motivations about which we are completely indifferent would be wanton and we would be non-autonomous (Frankfurt 1988: 164).

But here is where the *Regress/Arbitrariness Problem* looms. How does Frankfurt's hierarchical theory help identify which motivations are truly ours and which motivations arise from something outside ourselves? There must be some condition or requirement to help us determine when our motivations are *ours*. Frankfurt himself presents the *Regress/Arbitrariness Problem*: "Now it is pretty clear that this requirement cannot be satisfied simply by introducing another desire or volition at the next higher level. That would lead to a regress which it would be quite arbitrary to terminate at any particular point." (1988: 166). So how do we draw the relevant distinction between motivations? Frankfurt's answer is that motivations or desires with which we wholeheartedly identify are truly ours. The motivation or desire resounds within us; we feel like it would be pointless to search for further motivation. The motivation is decisive. We can stop at our second-order desire non-arbitrarily if the desire resonates within us and if we feel no conflict within ourselves regarding this desire. If we wholeheartedly endorse the second-

order desire, or if, in other words, we strongly identify with it, then no higher-order desires need be considered. Wholeheartedness is meant to provide a non-arbitrary stopping point for Frankfurt's hierarchical theory, thus avoiding the *Regress/Arbitrariness Problem*.<sup>23</sup>

### ***2.3. The Vagueness of Wholehearted Identification***

Unfortunately, the concept of wholehearted identification is vague. Is wholehearted identification merely synonymous with endorsement? At times it seems like this is what Frankfurt has in mind. But elsewhere, Frankfurt explicitly argues that identification is not equivalent to endorsement (2002: 161). Gary Watson proposes that identification, for Frankfurt, might be a way to express what one cares about (2002). But Frankfurt denies this understanding as well, noting that since he can identify with a desire to eat ice cream - a trivial desire - identification is not equivalent to caring (2002: 161). Wholehearted identification might also mean acknowledgement or acceptance. Frankfurt seems to adopt this definition when he argues: "The notion of identification that I do care about...is the notion of identification as acceptance" (2002: 161). But just before this passage, Frankfurt explains: "A person may be led to accept something about himself in resignation, as well as in approval or in recognition of its merit. The fact that he accepts it entails nothing, in other words, concerning what he thinks of it" (2002: 160).

I think what Frankfurt is trying to do, when clarifying the notion of wholehearted identification, is separate our identification with a desire from any positive feelings about the

---

<sup>23</sup> In his later work, wholeheartedness is also meant to explain the way in which a motivation or desire is central to an agent's identity. An agent is autonomous, on Frankfurt's most recent account, only when his "volitions derive from the essential character of his will". Motivations that reflect inessential parts of our characters are external to us, similar to the way in which motivations about which we are not at all concerned are external to us - that is, in an autonomy-undermining way (Velleman 2002: 92).

*content* of the desire. We need not think that our desires are *good* in order to identify with them. Identification, for Frankfurt, appears to be a sort of cognitive fit or coherence. It is primarily the avoidance of ambivalence about a given desire - an ambivalence that amounts to *incoherence* and *irrationality* (2002: 127). But I'm not sure if this clarification helps Frankfurt's account.

Recall that Frankfurt introduces the concept of wholehearted identification in order to avoid the *Regress/Arbitrariness Problem*. An agent wholeheartedly identifies with a certain desire if it resonates within her. Presumably, this resonance stops the regress because it signifies something about the agent - that the desire is properly *hers*, and that it matches *who she is*. But if identification has nothing to do with the positive feelings or endorsement of the agent, then it is unclear how it stops the regress in a non-arbitrary way. At best, then, the concept of wholehearted identification is vague; if, however, Frankfurt's later clarification properly narrows down the definition, it's no longer clear that the concept helps his account avoid the threat of a regress.

#### ***2.4. The Manipulation Problem***

There are several other objections raised against Frankfurt's account which reflect objections against hierarchical accounts in general. I will discuss two of these objections here.<sup>24</sup> The first objection, the *Manipulation Problem*, argues that because Frankfurt's theory is ahistorical, it cannot account for second-order desires that are placed in us through some sort of manipulation (Taylor 2005: 5; Christman 2007: 6). All that Frankfurt requires is that we wholeheartedly identify with our second-order volition that our first-order desire be effective. But we can easily

---

<sup>24</sup> Further objections include a metaphysical objection, which maintains that hierarchical theories like Frankfurt's presuppose a deep, or unified, or true self (Friedman, Oshana) and an agentic authority objection, which asks why a second-order volition should have the authority to speak for the self (Furrow and Wheeler 2013; Taylor 2005).

imagine a hypnotist giving us these second-order volitions *and* a feeling of wholehearted identification with these volitions (Christman 1989: 10, 2007: 6; Taylor 2005: 5, 8). In this case, our volition would have arisen from something *external* in exactly the way with which Frankfurt is concerned. But Frankfurt's theory requires that he conclude that the agent is autonomous in that case. This is not only counter-intuitive, but also seems to go against Frankfurt's own intuitions about external motivations.

### ***2.5. The Objectivity Problem***

The final objection I will rehearse here is the *Objectivity Problem*. Proponents of this objection - like T.M. Scanlon, Susan Wolf, and Barbara Herman - argue that Frankfurt's theory is problematically divorced from the objectivity of values.<sup>25</sup> Frankfurt's theory is purely procedural and internalist; in order to determine if an agent is autonomous, we need only examine her psychological desire structure. To determine if an agent has a particular reason for action, we need only look to her second-order volitions - we need not look at the *objective worth* of the object of her desire. But this focus on an agent's psychology, and accompanying lack of focus on objective worth, is problematic for theorists like Susan Wolf and Barbara Herman, who argue that a proper account of autonomy *must* include a reference to objective value. The objective value of an action must be taken into account when determining whether or not an agent has a reason for action; a mere consideration of her psychology is not enough.

On Susan Wolf's account, for example, what matters for free will is that we are capable of acting in accordance with Reason, where Reason refers to "the highest faculty or set of faculties there is, the faculty or set of faculties that, in most circumstances, will help us form *true*

---

<sup>25</sup> See their chapters in Buss and Lee Overton's 2002 volume *Contours of Agency: Essays on Themes from Harry Frankfurt*.

*beliefs and good values*” (Wolf 1990: 71, emphasis added). This means that the objective goodness of actions is built into the definition of free will, itself, which is why Wolf sees her account as more metaethical than other theories of free will (like Frankfurt’s) which are metaphysical (ibid.). Ultimately, acting in accordance with Reason amounts to “the ability to act in accordance with, and on the basis of, the True and the Good” (ibid.). To motivate her account, Wolf uses the example of someone who suffers through a depraved childhood. She considers JoJo, the son of an evil tyrant who has, as he has grown up, internalized his father’s sadistic desires (Wolf 1987: 54). Because JoJo (by stipulation) has so thoroughly internalized his father’s sadistic desires, and because he comes to endorse them as his own, Wolf argues that theorists of autonomy like Frankfurt would conclude that actions according to JoJo’s inherited sadistic desires are autonomous. Wolf disagrees with this analysis. Although JoJo may be able to reason, he is unable to act in accordance with right Reason because his values have been so distorted (Wolf 1990: 76). Free agency, in other words, requires the ability to track objective moral truth. Frankfurt’s theory fails because it does not give moral truth-tracking a central role in determining an agent’s autonomy.

The *Objectivity Problem* arises in response to Frankfurt’s lack of focus on objective value, while the *Manipulation Problem* arises in response to his lack of focus on an agent’s history. In the following two sections, I present two representative theories which attempt to avoid these problems. In the next section, I look at John Christman’s historical account, which is designed to avoid the *Manipulation Problem*, and in the following section I look at Gary Watson’s normative competency account, which is designed to avoid the *Objectivity Problem*.



### 3. Historical Theories

#### 3.1. John Christman

Like Frankfurt, Christman proposes an internalist, procedural account of autonomy. But unlike Frankfurt, Christman begins with two claims about persons: we are social beings, and we are diachronic beings. Selves are social in the sense that we are socially embedded; we always exist in a network of different relationships which help to form who we are. But for Christman, this is a contingent psychological claim about a person's values, emotions, or motivations rather than a metaphysical claim about what persons are as such (2004: 144). This distinction is important because it allows him to acknowledge the social nature of persons while maintaining the *individual* nature of autonomy (2004: 146).<sup>26</sup>

Selves are also diachronic - that is, they exist across time. It is important not only to acknowledge that people exist *socially* but also to acknowledge that they exist *historically* (2009: 96). Thus, for Christman, narrative coherence is required for autonomy. This is because, for Christman, “memory operates both as a constructive activity in giving us a sense of ourselves [...] but also in grounding current decisions as our own: such decisions cohere with the ongoing autobiographical narrative that grounds our agency” (Christman 2009: 138). Christman's motivation here is the same intuition that drives the proponents of the *Manipulation Problem* to take issue with Frankfurt's ahistorical, or 'time-slice' account: if an agent were to endorse her current values because of past manipulation or oppression (this will be especially important in the next chapter) our intuitions tell us that her endorsement is not really *hers* (2007: 6; 2009: 137). We want to say that she is missing something necessary for autonomy. But remember that on Frankfurt's account, an agent who endorses an action because a hypnotist hypnotized her into

---

<sup>26</sup> It is for this reason that I discuss him in this chapter, on traditional autonomy theories, instead of the next chapter on feminist relational theories.

this endorsement would 'count' as autonomous as long as she has the proper second-order volition. Christman's account avoids this counterintuitive result by examining persons as they exist *across time* instead of just *at a time*.

Another important difference between Frankfurt and Christman is that, while Frankfurt is concerned with the autonomy of an agent regarding specific choices, Christman is concerned with the autonomy of an agent regarding her basic orienting values or motivations (2009: 136). So while it made sense to talk of Betsy and her ice cream using Frankfurt's rubric, this scenario would almost certainly not be something with which Christman is concerned. Christman's account thus falls somewhere in the middle between strictly local accounts of autonomy on the one hand, and strictly global accounts on the other.

Broadly speaking, Christman argues that there are two types of conditions that are jointly sufficient for autonomy.<sup>27</sup> The first are *competence conditions* concerning an agent's rationality and self-control, and the second are *authenticity conditions*, concerning whether or not an agent's values represent who she is as a person (2007: 3-4). Specifically, authenticity conditions determine whether or not a specific value fits into a coherent autobiographical narrative for the agent (2009: 137).<sup>28</sup> These two types of conditions make up Christman's definition of autonomy (2009: 155):

An agent A is autonomous regarding a basic orienting value or motivation V if:

[I. Basic Requirements - Competence]

1. A is competent to effectively form intentions to act on V;
2. A has the general capacity to critically reflect on V;

[II. Hypothetical Reflection Conditions - Authenticity]

---

<sup>27</sup> Importantly, Christman argues that his conditions are merely sufficient for autonomy, but not necessary. That is, there may be a different set of conditions that would similarly render an agent autonomous.

<sup>28</sup> Note that it is an *autobiographical* narrative, since Christman's account is subjective (internalist).

3. Were A to engage in critical reflection on V in light of the historical process (adequately defined) that gave rise to V,
4. A would not be alienated from V, and
5. The reflection (even if hypothetical) is not constrained by reflection-distorting factors like manipulation.

The competence conditions (1-2) are fairly straightforward and map onto Frankfurt's Desire and Volition conditions, but the authenticity conditions require more explanation. The first thing to note is that Christman requires only *hypothetical* and piecemeal reflection on a value rather than actual reflection (2009: 145). Christman argues that many of the factors that make us who we are as persons are never chosen by us, and might similarly never be consciously accepted by us (2009: 145); but this doesn't mean that they don't make up an essential part of our self-conception. Christman argues: "Many attitudes and motives that we are guided by and express our settled sense of ourselves were never actually chosen by us, though in being so guided we accept them without conflict (upon reflection)" (2009: 146). The important thing, for Christman, is that the values or motivations in question fit coherently into our historical narrative, or would fit coherently into our historical narrative were we to reflect upon them.

The second thing to note about Christman's authenticity conditions is that determining whether or not a value fits coherently into our historical narrative is not a matter of determining whether or not we identify with it. Identification of the sort required by Frankfurt is too strong; any theory that requires identification with our values, let alone wholehearted identification, would dictate that "only the supremely lucky and fulfilled among us would count as autonomous" (2009: 13). Most of us (or as Christman says, "the rest of us sorry sort") often feel ambivalent or indifferent about a given value without fully rejecting it. What should be required for autonomy, then, instead of identification, is non-alienation. Alienation, for Christman, is both cognitive and affective; to be alienated is to feel a strong rejection, which might present

itself as a negative affect (2009: 144). If we feel alienated from a given value, if we feel constrained by it or if we want to rid ourselves of it, but can't, the alienation will render us non-autonomous. The final thing to note about Christman's account is that he stipulates that the critical reflection necessary for autonomy must not be constrained by reflection distorting factors like manipulation or oppressive upbringing.

### ***3.2. The Agentic Authority Problem***

There are several objections that can be raised against Christman's account. I will discuss two here: the *Agentic Authority Problem* and the *Limited Subjectivity Problem*. The *Agentic Authority Problem* arises out of the concern that our reflective endorsement may be inauthentic. Christman is right to stipulate that the reflection (hypothetical as it is) must not be the product of manipulation, since such manipulation may not adequately reflect the identity of the agent. But since it is possible that someone could be manipulated into reflectively endorsing a given value, how are we to be sure that our endorsement has not been manipulated? Christman's answer is that the reflective endorsement must be repeated over a variety of circumstances (2009: 152). If the reflection were repeated several different times, and if the result of the endorsement resulting from the reflection continued to be the same, we could conclude that the reflection was authentic. Christman's account thus requires "sustained critical reflection" or SCR. Christman explains: "The idea is that when a person reflects on a trait over time and in a variety of settings and contexts, always yielding neither alienation nor rejection, such reflection indicates the kind of settled character that autonomous agency manifests" (2009: 153).

The problem with SCR as a proposed response to the *Agentic Authority Problem* is that it doesn't adequately deal with the problem it is designed to solve. If it is possible for a hypnotist,

for example, to hypnotize us into endorsing a given value, then it is also possible for that hypnotist to hypnotize us into endorsing it over a long period of time and in a variety of circumstances. Christman might respond that we would have to compare the endorsement at time  $t_2$ , after the agent was hypnotized, with the endorsement at time  $t_1$ , before the agent was hypnotized. If these two endorsements cohered, we could conclude that the agent was autonomous relative to the desire in question. But while this solution might work in the example of a hypnotist, it is less likely to work in cases involving oppressive upbringings, which Christman also claims SCR can address. It isn't possible to compare an agent's endorsement at time  $t_2$ , after an oppressive upbringing, to her endorsement at time  $t_1$ , before her upbringing. It's not clear, then, that SCR adequately avoids the *Agentic Authority Problem*.

### ***3.3. The Limited Subjectivity Problem***

The second objection to Christman's account that I will discuss is the *Limited Subjectivity Problem*. Christman argues that his theory is entirely subjective, that is, it determines autonomy based on the psychology of the agent and not on external factors. He presents several arguments against non-subjective, or external, theories to motivate the subjective nature of his own theory - including an argument that subjective accounts allow different agents to have different types of value-development while maintaining their autonomy. But objectivity creeps in at several places throughout his discussion, most notably in his proposition of hypothetical reflection. Consider his example: a woman finds out that she had been abused as a child by someone who was responsible for many of her current skills. After her memories of the abuse surface, she repudiates many of these skills. She doesn't remember, however, that her abuser also taught her to play the piano, and so she does not repudiate this specific skill. If she were to be told that her

piano-playing skill came from her abuser, she would feel alienated from that part of herself. Is this person autonomous? Christman says *no*, because *were* she to reflect on her skill in light of its origins, she would be alienated from it (2009: 157-8). Presumably, this is why he includes the parenthetical in condition 3 above, that the history behind the skill must be adequately defined.

But the adequate definition of the history of the agent's piano-playing skills, in this case, is a definition that is not subjective. Instead, it relies upon an objective, or fact-of-the-matter description that, in this case, is unknown to the agent. The autonomy that results from the agent's reflection in this case - because the reflection is merely hypothetical - doesn't have to belong to the agent at all. In fact, in this case, the reflection on the actual origin of her skill cannot belong to the agent, because she does not possess the necessary information. But how is this meant to be part of a subjective theory? Christman argues: "If people are under illusions about their own character - illusions that would cause them severe internal distress leading to self-repudiation if made clear - they are not really autonomous relative to them" (2009: 158). But this is to impose external constraints on an agent's autonomy. Christman's account does not, then, look fully subjective.<sup>29</sup>

## **4. Normative Competency Theories**

### ***4.1. Gary Watson***

So far, the two theories considered have been procedural and internalist; they have been concerned with the psychological coherence of an agent's desire structure and with her endorsement of her desires. Gary Watson, however, argues that an agent's endorsement is not enough to ensure free agency, which he defines broadly as the ability to do or get what one wants

---

<sup>29</sup> Note that this is not an objection against Christman's view *per se*, but rather against his classification of his view as fully subjective.

(1975: 205). Endorsement is sufficient for mere intentional action, but intentional action and autonomous action are not the same. Autonomous action requires something more than intentional action - it requires an appeal to our normative values.

Watson makes a crucial distinction that plays a central role in his theory of autonomy, between wanting (or desiring) and valuing. Up until this point, I have been using desire and value interchangeably, and this reflects the way that many autonomy theorists (including Frankfurt and Christman) use the terms. But according to Watson, wanting X and valuing X cannot be used interchangeably because they are not equivalent. If they were equivalent, then it would have to be the case that they would always coexist, so that if an agent wanted X she also valued X, and if she valued X she also wanted X. But Watson argues that this is simply not the case. It may be the case, for example, that one does not value at all what one desires (1975: 209-10); an agent may desire X without thinking that it is worthwhile or good. It might also be the case that an agent could strongly desire something while valuing it to a lesser degree (1975: 210). In other words, when an agent values X, she judges that X is good; when an agent desires X, she has not necessarily judged anything - the desire might simply be an expression of her appetites or passions (1975: 212). He asks us to consider the case of a squash player who has suffered "an ignominious defeat" (1975: 210), and momentarily desires to smash in the face of his opponent with his racket. Watson argues that it is simply *false* that the squash player values the smashing in of his opponent's face, even though he desires it; and these sorts of cases do not just occur in cases of momentary or fleeting desires.

Watson explains that this distinction is not based on the differing *content* between the value and the desire; instead the distinction regards the source of the desire and the role that the desire plays in the "total 'system' of the agent's desires and ends" (1975: 211). Because values

and desires have different sources, they make up different systems within an agent. An agent's valuational system is the set of considerations that yield judgments such as: All things considered, Z is the best thing for me to do in these circumstances (1975: 215). An agent's motivational system, on the other hand, is the set of considerations that move her to act. That these two can come apart can be shown using the Betsy example used earlier. Betsy might have the judgment that, all things considered, not eating the ice cream is the best thing for her to do while still having a motivating desire to eat the ice cream. In this case, Watson would argue that when Betsy eats the ice cream, she is not getting what she really wants (because she has done something that she judged to be bad) and so is non-autonomous.

But importantly, for Watson, this non-autonomy is not the result of an incoherence between a second-order volition and a first-order desire. The lack of Betsy's endorsement is not what makes her non-autonomous regarding her choice to eat the ice cream. Instead, Watson argues that any action which is in accord with an agent's evaluational system is an action which expresses what an agent really wants, and thus can be said to truly belong to the agent. Consider another example. Greg is an avid hiker who has recently suffered a minor ankle injury. His doctor has instructed him to stay off his ankle, and thus avoid hiking for three weeks in order to give his ankle time to heal. Greg values his good health, and so judges that, all things considered, staying off his ankle for three weeks is the best thing for him to do under the circumstances. But after a week, he becomes bored. The weather is beautiful and he can no longer stay inside. Even though he values his good health, and even though he has judged that he should obey his doctor, his passion for hiking wins out, and Greg goes on a hike.

In this example, Greg's self-control has failed him. He had an evaluational system that disclaimed hiking after only a week, but a motivational system that moved him to hike anyway.



In this case, Watson would argue that Greg is not getting what he really wants, or in other words, Greg's actions are not truly his own (1975: 219). Remember that Watson's definition of autonomous action is an action that is an agent's own, in the sense of being what an agent truly wants - an agent acts autonomously to the extent that the action in question is an expression of what he really values. Watson argues: "What is distinctive about such compulsive behavior [like Greg's decision to hike] ...is that the desires and emotions in question are more or less radically independent of the evaluational systems of [the] agents" (1975: 220). Since the evaluational system is what is most definitive of an agent's identity, that is, because an agent cannot separate herself from her evaluational system, any action that does not cohere with her values is an example of a non-autonomous action. Our actions can only be autonomous, then, when they are consistent with what we value as good, and our values must be the right ones.

#### ***4.2 The Influence of Desire Problem***

Unfortunately, Watson's account suffers from some of the same weaknesses as Frankfurt's - it is unable to account for values that come from a source other than the agent herself. In other words, Watson's account does not provide an adequate response to the *Manipulation Problem*. Just as with Frankfurt's account, it would be possible for a hypnotist to instill certain values in us that we would come to feel were definitive of our identities. Watson's account thus suffers from the same ahistorical nature from which Frankfurt's account suffers. There is still much more to say about this complicated topic, and I will say more in the next chapter. For now I will raise one final objection against Watson's theory - the *Influence of Desire Problem*.

According to Watson, a person's values are what most accurately reflect who that person is. We are never able to tear ourselves away from our evaluational system, and for this reason it

defines our identity. But what Watson fails to take into account is the way in which our desires or wants might actually shape our evaluational system. Consider a child who grows up in a poor household, unsure of where his next meal will come from. This child has a constant desire for food. But there is a sense in which this consistent desire shapes his values - he does not only *desire* food, he *values* it as a good. Watson does make the following claim: "A [...] possibility that presents considerable problems for the understanding of free agency is this: some desires, when they arise, may 'color' or influence what appear to be the agent's evaluations, but only temporarily. That is, when and only when he has the desire is he inclined to think or say that what is desired or wanted is worth while or good" (1975: 214; cf. Buss 2012: 658n.25). But a desire as strong as hunger might influence our values permanently, not just for the moment we feel it the most.

Note that the problem here is not the consistency, or coherence, between the desire for food and the potential evaluation of food as a good. This coherence would, for Watson, signal the child's autonomy. The problem is that what is foundational within the child is not his evaluational system, but rather his desire for food. It is his hunger that motivates his values; his desire for food does not flow from an independent value set. The *Influence of Desire Problem* illustrates the complex relationship between values and desires - a complexity that is not fully appreciated by Watson. In the next chapter I turn to consider feminist objections to traditional theories of autonomy.

## CHAPTER III

### RECENT WORK ON RELATIONAL AUTONOMY

Recent work on relational autonomy can be bookended by two pieces: Catriona Mackenzie and Natalie Stoljar's landmark volume, *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (2000), and Andrea Veltman and Mark Piper's recent collection, *Autonomy, Oppression, and Gender* (2014).<sup>30</sup> Veltman and Piper's collection began as a conference, *Relational Autonomy: Ten Years On*, held to reassess the state of the field since Mackenzie and Stoljar's pivotal volume. It is therefore an appropriate representation of the maturation of the debate. In this chapter, I review the recent literature on relational autonomy and highlight the changes in the field in the past fifteen years. I conclude by suggesting that dichotomies which once defined the debate are beginning to break down, and are being replaced by nuance at the theoretical and meta-theoretical level.

#### **1. Relational Autonomy: The Early Years**

##### ***1.1. The Self: Unified, Transparent, Rational, Individual?***

At the time of the release of Mackenzie and Stoljar's *Relational Autonomy* (from here on, RA) the concept of autonomy was still under fire from many feminist theorists for its overly individualistic and rationalistic nature.<sup>31</sup> In the Introduction to RA, Mackenzie and Stoljar helpfully identify five major types of feminist critique, all of which reject the conception of the self and the conception of autonomy implicit in mainstream theories (Mackenzie and Stoljar

---

<sup>30</sup> This is of course not to downplay earlier feminist work on this topic. See e.g. Meyers 1987, 1989; Nedelsky 1989; Benson 1990, 1991; Govier 1993; Friedman 1997.

<sup>31</sup> See e.g. Jaggar 1983; Code 1991.

2000: 5-12). Feminists were worried about mainstream theories' autonomous self. Postmodernist feminists, for example, challenged the idea of a transparent, unified self allegedly at play in traditional theories of autonomy.<sup>32</sup> John Christman's idea, for example, that we must be able to reflect upon and endorse (or at least not be alienated from) our values and motivations requires that we be able to know what our values and motivations are. In fact, all internalist theories of autonomy encounter this problem, since they appeal to our internal psychological processes or states to determine whether or not we are autonomous. The reasoning behind the endorsement requirement is that such endorsement will reflect who we *really are*; if we endorse a given value, that value must represent our *true selves*. Recall Frankfurt's argument that we must wholeheartedly identify with a second-order desire in order for that desire to be autonomous. Wholehearted identification ensures that the desire in question is truly *ours*. Autonomous action is authentic action.

But postmodernist feminists, poststructuralist feminists, and feminists influenced by psychoanalytic theory have argued that authenticity conditions or endorsement requirements presuppose a transparent and unified self. There are two parts to this challenge: an epistemological thesis - that we are not self-transparent - and a metaphysical thesis - that selves are intersectional instead of unitary (Scheman 1993: 96; Lugones 1997: 154; Code 2000: 182; Meyers 2000: 153, 2010: 5-6; Brison 2005: 369; Grimshaw 2005: 337). Feminists influenced by psychoanalytic theory use the Freudian concept of the unconscious to demonstrate that we are not always able to know ourselves. Our unconscious may hide from view aspects of our

---

<sup>32</sup> As Mackenzie and Stoljar point out, it's not clear that there are any postmodernist critiques of traditional autonomy as such; rather, they critique the model of selfhood that is implicit in these accounts. See Mackenzie and Stoljar 2000: 10.

development as persons, or it may make it difficult for us to make sense of conscious experiences (Grimshaw 2005: 333).

It is important to note that this critique is not simply a charge that traditional autonomy theories have gotten something wrong about the nature of the self. Diana Meyers points out that Frankfurt-type accounts of autonomy which require wholehearted identification are not adequate for understanding oppression or victimhood. A person might wholeheartedly disidentify with her oppression based on her group membership, while still identifying with the group membership itself. This sort of ambivalence is non-autonomous on Frankfurt's account, leading Meyers to conclude that wholehearted identification is an inadequate criterion for intersectional selves (Meyers 2000: 170).

In addition to criticizing the idea of a unified and transparent self, some feminists have also criticized the idea of a rational self. More specifically, critics charge that reason, especially the concept of reason arising from the Enlightenment, has been associated with masculinity (for discussion, see e.g. Nagl-Docekal 1999; Heikes 2010). The defining features of Enlightenment rationality are features that have historically been ascribed to men: reason is disembodied, always conscious, and objective. The opposites of each of these – embodiment, the unconscious or emotional, and subjectivity – are traits historically ascribed to women. Disembodied reason, for example, is an ideal that rests on a Cartesian dualism between mind and body. The idea that rationality must always be conscious rests on a dichotomy between reason and emotion, while the ideal of objectivity rests on a dichotomy between objective truth and subjective particularity. Historically, each of these dichotomies has had a 'good' or 'right' element and a 'bad' or 'wrong' element, where the 'good' side (disembodiment, for example) is understood as masculine and the 'bad' side (embodiment) is understood as feminine.

The problem with these dichotomies is not simply that they are descriptively wrong; each dichotomy also contains a normative judgment about what sorts of knowledge are real – a phenomenon that Alison Jaggar dubs “normative dualism” (Jaggar 1983). The ideal of objectivity, for example, requires that proper reason must meet certain standards or rules, and these rules apply to all people (Longino 2005: 80). Thus, contextual or situated reason is not *properly reason*. Some feminists, especially postmodern feminists, have argued that this normative judgment decides against women, deeming them unreasonable or irrational and thereby dismissing them as knowers and silencing them. Helen Longino presents the postmodernist feminist worry: “What the postmodernist and the feminist resist is the idea that there is a template of rationality in which all discourses fit, a template that dissolves the barriers of locality...Such a template or universal language would...impose a straightjacket of uniformity disallowing the diversity produced by differently situated perspectives. One person’s rationality, then, is another’s tyranny” (Longino 2005: 81).

While these latter two critiques are both common and well-known, the central critique of the theory of the self at play in traditional autonomy theories is focused on a problematic individualism. It is to this issue that Mackenzie and Stoljar devote a majority of their introduction, and I would argue that it is on this issue that much of the relational autonomy debate (especially the early debate) has been focused. The most “entrenched” of these critiques of individualism is the metaphysical critique (Mackenzie and Stoljar 2000: 7), which charges mainstream theories of autonomy with presupposing an individualistic conception of the self. This charge, in turn, can be broken down into four different objections, which charge mainstream theories with presupposing the following: first, that agents are causally isolated from others, when in reality we are all – at some time or another – dependent (Baier 1981); second, that

agents' *self-conceptions* don't rely on any others, when in reality our self-conceptions are constituted by our social contexts; third, that *agents* are not constituted at all by their social relations, when in reality we are; and finally, that (metaphysically speaking) agents are separable beings (Mackenzie and Stoljar 2000: 7).

The last of these claims is easily dismissed as an objection – there are no theories of autonomy that reject its truth. But the first three objections linger. Jennifer Nedelsky's important 1989 piece, "Reconceiving Autonomy: Sources, Thoughts and Possibilities", is one of the most forceful representations of the metaphysical critique.<sup>33</sup> Contrary to the "liberal vision" at the heart of mainstream autonomy theories, that selves are atomistic individuals, Nedelsky argues, with feminists and communitarians, that selves are constituted by their social relations (Nedelsky 1989: 8). Mainstream theories of autonomy, and the liberal individualism at their centers, reject what women experience as reality: we are all socially embedded. Unless autonomy can be reconceived, it will continue to fail to represent the real lives of men and women; it will continue to be a false and oppressive ideal.

### ***1.2. Autonomy: Isolationist, Individualistic, Masculine?***

Far from rejecting the concept of autonomy, however, Nedelsky recognizes that it is required to meet feminist goals. Feminists "need a language of freedom" so that women can "define who we (each) are, rather than accepting the definition given to us by others (men and male-dominated society, in particular" (Nedelsky 1989: 8-9). Nedelsky thus acknowledges a distinction between *theories* of autonomy and the *concept* of autonomy as such. In this section, I

---

<sup>33</sup> Although Mackenzie and Stoljar (I think mistakenly) do not cite her in their discussion of the metaphysical critique, they do give her credit for being the first to advocate a relational theory of autonomy (Mackenzie and Stoljar 2000: 26n.1).

will briefly mention early feminist critiques of the *concept* of autonomy before moving on to discuss the major objections to mainstream *theories* of autonomy.

RA's main goal was to salvage the concept of autonomy from earlier critiques which held that the concept of autonomy itself was descriptively and prescriptively flawed. Descriptively, autonomy was seen as a nebulous ideal that didn't reflect the lives of women. It was thought that, to be an autonomous agent, one had to eschew the influence of others and free oneself from social relationships in order to become an independent, self-made man. Such an ideal may fit the lives of men, who have real-world Autonomy Icons like Henry David Thoreau, Ralph Waldo Emerson, and Christopher McCandless, and fictional Autonomy Heroes like James Bond, Clint Eastwood's 'The Man with No Name', and (more recently) Christian Bale's Batman and Matt Damon's Jason Bourne (to name but a few). One is hard pressed to find so many female Icons or Heroes; Emily Dickinson, who immediately springs to mind, is not praised like Thoreau, but is instead either mocked or pitied for her recluse nature.<sup>34</sup> The concept of autonomy was also thought to be prescriptively flawed, since it promotes values like social isolation and self-sufficiency.

There are two things to note about these earlier feminist critiques. First, Marilyn Friedman convincingly argues that these Icons and Heroes represent a colloquial understanding of autonomy as rugged independence, rather than a philosophically sophisticated conception (Friedman 2003).<sup>35</sup> Second, by the time RA was published, the push for a feminist salvage mission was at its peak. Many of the contributors thus take a two-pronged approach: distinguish

---

<sup>34</sup> Even her closest friend, Thomas Wentworth Higginson, is known to have described her as "partially cracked".

<sup>35</sup> She specifically mentions Paul Gauguin and the Marlboro Man.



theories of autonomy from the concept of autonomy itself, and then demonstrate that, while the former might be faulty, the latter has enormous feminist potential.

The objections to mainstream theories can be broken down into two broad types: objections to what the theories say, and objections to what they fail to mention. Here I will discuss feminist objections to Harry Frankfurt and John Christman, since these are (arguably) the most prevalent. Recall the description of Frankfurt's hierarchical theory from the previous chapter: according to Frankfurt's initial 1971 theory, an agent must meet two requirements in order to be autonomous. First, she must have a second-order desire that supports her first-order desire. If Agent A has an effective first-order desire to  $\phi$  - that is, she *wills*  $\phi$  - she must have a second-order desire to want to desire  $\phi$  - that is, she must have a second-order desire that she will  $\phi$ . But Agent A must also want her first-order desire to move her to action; she must want her first-order desire to  $\phi$  to be effective (Taylor 2005: 4). Using Frankfurt's terminology, she must have a second-order *volition* that her first-order desire move her to  $\phi$ . In 1987, Frankfurt amended his theory to include the wholeheartedness condition - in order for a second-order desire to be autonomous, an agent must wholeheartedly identify with it.

The primary feminist objection to Frankfurt's theory is that it is wholly psychological. Although the introduction of the wholeheartedness condition helps to alleviate the *Regress/Arbitrariness Problem*, it does nothing to take into account the effects that external factors have on our internal psychologies. It is entirely possible on Frankfurt's hierarchical account that an agent wholeheartedly identify with a second-order volition that a first-order desire be effective, when both the *desire* and the *identification* were implanted in her through some form of manipulation (Taylor 2005: 5; Christman 2007: 6). This is the *Manipulation Problem* discussed in Chapter Two. But this problem is especially poignant for feminists, who

are explicitly concerned with the problem of internalized oppression. It is possible, on Frankfurt's account, that a woman wholeheartedly identify with an oppressive desire (for unquestioning submission to her husband, for example) even though the initial desire was 'implanted' in her by a patriarchal society. Paul Benson argues that, "Since the socialization of feminine appearance is most effective when norms of femininity are securely internalized in women's attitudes towards themselves, women's highest-order desires can fall under its sway just as completely as can first-order desires" (Benson 1991: 391). The wholehearted identification is due to the fact that the agent has internalized patriarchal norms as her own, and thus serves to further her own oppression.<sup>36</sup>

Diana Meyers provides a more nuanced critique of Frankfurt's wholeheartedness condition (Meyers 2000: 170-2). Because people are intersectional, rather than unitary, subjects, it is unlikely that we will be able to identify wholeheartedly with any second-order volition. We are likely to feel ambivalent about some of "identity-determinants" (Meyers 2000: 170), while wholly embracing others, and decidedly disidentifying with others; in other words, we are bound to feel tension between our various identities. But this tension, argues Meyers, is actually the source of necessary social critique (Meyers 2000: 170); thus, we should not try to eliminate all tension or ambivalence. This observation is squarely opposed to Frankfurt's wholeheartedness condition. But Meyers also makes a further point: because our intersectional identities are really ours, and because these identities may often conflict or be in tension with one another, to try and eliminate all conflict or tension in the name of wholehearted identification is to *deny who we are*. But the purpose of the wholeheartedness condition in the first place was to ensure authentic agency; Frankfurt thus fails at his own task (Meyers 2000: 170).

---

<sup>36</sup> For an extended discussion of internalized oppression, see Chapter Four.

Feminists have also criticized John Christman's historical account.<sup>37</sup> Recall Christman's definition of autonomy from Chapter Two (Christman 2009: 155):

An agent A is autonomous regarding a basic orienting value or motivation V if:

[I. Basic Requirements - Competence]

1. A is competent to effectively form intentions to act on V;
2. A has the general capacity to critically reflect on V;

[II. Hypothetical Reflection Conditions - Authenticity]

3. Were A to engage in critical reflection on V in light of the historical process (adequately defined) that gave rise to V,
4. A would not be alienated from V, and
5. The reflection (even if hypothetical) is not constrained by reflection-distorting factors like manipulation.

Feminist criticisms of Christman's account focus on (3) and (4) above – the hypothetical endorsement criteria (or, better, the hypothetical lack-of-alienation criteria). For Christman, the desire in question must fit coherently into our historical narrative, or *would* fit into this historical narrative *if* we were to reflect on it. But feminists have argued that the hypothetical endorsement criteria do not do enough to weed out internalized oppression. Paul Benson, for example, points out that a thoroughly socialized agent would likely not revise an oppressive desire, even if she were to reflect on it (Benson 1991: 392-4; cf. Mackenzie and Stoljar 2000: 16). This might be the case because even the process of reflection itself would be influenced by oppressive socialization. Benson argues, for example, that a woman who desires to look feminine could have “become accustomed to thinking of herself from an internalized male point of view, so she may be unaffected by the knowledge that her endorsement of her desire to have a feminine

---

<sup>37</sup> In the previous chapter I expressed some hesitation about including Christman as a ‘mainstream’ theorist instead of a feminist theorist. Because he is treated as a mainstream theorist in much of the feminist literature, I continue to categorize him in this way. But I think there is a good argument to be made in favor of considering his account a rival *feminist* theory.

appearance was the product of socialization in a male-dominated society” (Benson 1991: 394). Internalized oppression runs so deep that the process of endorsement may itself fall prey.

In short, feminist objections to both Frankfurt and Christman focus on the fact that their theories cannot adequately take account of internalized oppression. Their theories fail in this way, it is argued, because they do not acknowledge the social nature of the self and the social nature of autonomy. But again, this is a problem with theories of autonomy, and not with the concept of autonomy itself.

### ***1.3. Reclaiming Autonomy***

This distinction also forms the basis of Mackenzie and Stoljar’s defense of autonomy for feminist purposes. Although mainstream theories may fail to take into account the social nature of persons, there is nothing about autonomy as such which precludes such an understanding.

But autonomy also has a positive value to bring to feminist theory. Demonstrating the importance of autonomy for achieving feminist goals was a central component of the feminist reclamation of the concept of autonomy. Marilyn Friedman’s chapter in RA, “Autonomy, Social Disruption, and Women”, goes a great distance toward assuaging what she calls “autonomophobia” or the fear of autonomy (Friedman 2000: 35-51). Autonomophobia, Friedman explains, is a legitimate worry – especially for women, who often depend on social relationships for “material and emotional sustenance” (Friedman 2000: 45). But despite this dependence – indeed, *because* of it – women are sometimes oppressed and abused by the very people on whom they depend. So rather than fearing autonomy because it may lead to the disruption of some social relationships, women should welcome it as a cultural ideal; the

relationships disrupted may in fact be harmful. Not all social relationships are valuable; not all autonomy is dangerous (Friedman 2000: 47).

This theme is echoed throughout RA. Diana Meyers, for instance, argues against thinking of autonomy as an “androcentric phantasm” (Meyers 2000: 152). Feminists have rightly objected to the “self-originating, self-sufficient, coldly rational, shrewdly calculating, self-interest maximizing, male paragon of autonomy”, but this is a caricature (Meyers 2000: 152). To the contrary, she argues that autonomy should be a value of great importance to feminists, since it is autonomy that forms at least part of the story about women who escape oppressive relationships, protest against patriarchal norms, and define themselves as persons even while dominated and subordinated. Of the contributions to the volume that don’t explicitly argue for the importance of autonomy, almost all implicitly reflect it.

## **2. Procedural or Substantive: Never the ‘Twain Shall Meet**

### ***2.1. From Procedural to Substantive***

After the publication of RA, there was a noticeable shift in the relational autonomy literature. Whereas most feminist theories of autonomy up until this point had been procedural (Friedman 2003; Meyers 1989), more and more feminists began adopting substantive theories. This is likely due, at least in part, to Natalie Stoljar’s influential piece in RA, “Autonomy and the Feminist Intuition” (Stoljar 2000: 94-111). In this section I will briefly explain the strengths of procedural theories before moving on to discuss the shift to substantive theories.

Procedural theories of autonomy are content-neutral, that is, there are no direct restrictions on the *content* of an autonomous person’s beliefs or desires. This content-neutrality made procedural theories especially attractive to early relational autonomy theories, because it

provided a buffer against feminist concerns with autonomy as a masculine, individualistic concept. If there is no direct restriction on an agent's desires, she could potentially endorse any desire – including a desire for close relationships – without thereby being deemed non-autonomous. Procedural accounts thus allow for the coexistence of autonomy and relational desires; rather than requiring robust individualism and rugged independence, procedural theories allow autonomous agents to value their relationships. But the sword cuts both ways; procedural accounts place no direct restrictions on the content of an agent's desires, and so do not (based on this information alone) discard as non-autonomous problematic desires (like a desire for subservience, for example).<sup>38</sup> An agent with a desire to submit to her husband could *still be autonomous* if she has reflectively endorsed this desire (Friedman 2003: 5).

Diana Meyers' has a procedural account of autonomy according to which an individual can achieve autonomy only by participating in processes of self-discovery, self-definition, and self-direction (Meyers 2004: 16).<sup>39</sup> The specific skills that make up *agentic competency* or *autonomy competency* include skills of introspection, communication, memory, imagination, reasoning, self-nurturing, and volition (Meyers 2014: 121). The extent to which agents engage in these skills is the extent to which they are autonomous.<sup>40</sup> Like Friedman's account, the procedural nature of Meyers' account allows her to treat women with desires for submission as not necessarily non-autonomous. It's important to note, for both Friedman and Meyers, that the

---

<sup>38</sup> Friedman considers the example of a woman who chooses to remain in an abusive relationship. If the desire to remain authentically belongs to the agent, then it is at least possible that she chooses autonomously (Friedman 2005: 146).

<sup>39</sup> Meyers concedes in 2008 that her account is weakly substantive, rather than procedural: "Paul Benson has convinced me that there is a spectrum of positions between the poles of strict value-neutrality and strong value-commitment and that, however different our views are in some respects, he and I are currently advocates of 'weak substantive' positions" (Meyers 2008: 206). In her most recent work, however, Meyers reclassifies her theory as value-neutral, thus reversing her concession to Benson (Meyers 2014: 121n.19).

<sup>40</sup> Since, for Meyers, autonomy admits of degrees (Meyers 2004: 9).

procedure involved in an agent's decision-making process can still weed out oppressive preferences. For Meyers, oppressive socialization can threaten an agent's agentic competency and her self-respect (Meyers 1989: 208),<sup>41</sup> and Friedman acknowledges that oppressive relationships can have a huge impact on the way that an agent reflects on her desires (Friedman 2003: 97).

But other feminists are not convinced that procedural theories could take account of internalized oppression. In "Autonomy and the Feminist Intuition" Natalie Stoljar argues that neither content-neutral nor weak substantive accounts of autonomy are able to uphold the feminist intuition, which claims that, "preferences influenced by oppressive norms of femininity cannot be autonomous" (Stoljar 2000: 95).<sup>42</sup> Stoljar utilizes Kristen Luker's study of women who engaged in contraceptive risk taking as an illustration of internalized oppression (Stoljar 2000: 99). Autonomy theories which only examine factors internal to the agent cannot separate oppressive preferences from legitimate preferences, and so cannot conclude that Luker's women are non-autonomous – even though Stoljar claims that they clearly incite our feminist intuition. Only a strong substantive theory is able to account for the heteronomy of Luker's women, because only a strong substantive theory examines the *content* of the norms they endorse; it is this type of theory that Stoljar ultimately endorses.<sup>43</sup>

Part of the reason why Stoljar's paper was so influential is because it successfully switched the burden of proof onto procedural theorists. Anyone who wished to endorse a

---

<sup>41</sup> It is this feature of Meyers' account that leads some to conclude that her account is weakly substantive.

<sup>42</sup> Thus, whether Meyers' account is classified as procedural or weakly substantive, Stoljar's critique is still meant to apply.

<sup>43</sup> She refers to her own theory as a strong substantive theory in her 2000 piece, but at times her description of her own account makes it sound like a normative competency account (see my Chapter 4). She later classifies her account as a normative competency account, correcting her earlier self-classification (Stoljar 2014: 229n.10).

procedural account had to answer Stoljar's question: how can you uphold the feminist intuition? Rhetorically, this question implies that there is a choice to be made – between *procedural* theories, on the one hand, and *feminist* theories on the other. And following the uptake of RA, many feminist autonomy theorists started opting for the latter (Benson 2005; Superson 2005; Oshana 2006; Mackenzie 2008; Charles 2010; Stoljar 2014). Of course, substantive considerations can be taken into account in different ways (for example, normative competence theories) and to different degrees (strong or weak). What is important for my purposes here is to note the shift that occurred following the enormous influence of Mackenzie and Stoljar's 2000 volume.

## ***2.2. Autonomy as Social-Relational***

So far I have been discussing procedural and substantive theories without direct reference to relationality. What is it about theories like Stoljar's that make them *relational* autonomy theories? Why does Friedman's account, which pays great attention to the effects of socialization, often '*not count*' as relational? What is the relationship between the degree of substantiveness of a theory and the degree of relationality? First, a distinction: theories of autonomy can either be causally relational or constitutively relational. If they are *causally* relational, they recognize the impact that social forces have on agents and their autonomy. If they are *constitutively* relational, on the other hand, they think that these social forces actually constitute what it means to be autonomous. Procedural theories are (most often) causally relational, while substantive theories are (most often) constitutively relational.<sup>44</sup> Procedural theories, which are only causally relational, are thus not uniquely relational – this leads some to

---

<sup>44</sup> I will discuss an exception, dialogical theories, later.



classify Friedman's (and Christman's) accounts as not relational (e.g. Benson 2014: 96-104). Relational theories of autonomy are those that are uniquely relational; this means that, definitionally, relational theories of autonomy are going to be *constitutively* relational.

On the spectrum of procedural to substantive accounts, strong substantive accounts like Stoljar's give the most weight to social relationships. Marina Oshana also endorses a strong substantive account, what she calls a social-relational account of autonomy, which requires that, in order to be autonomous, an individual must be in certain sorts of relationships, and that she must have chosen to be in these relationships. On Oshana's account, it can be objectively determined which sorts of relationships are conducive to or destructive of autonomy.<sup>45</sup> However, an individual might reflectively endorse a given relationship which is objectively found to be a hindrance to her autonomy. In this case, an individual's own assessment of her autonomy can conflict with the *matter of fact* (Oshana 2005: 91-3). Therefore, on Oshana's account, an agent's autonomy could be undermined based on the type of relationships of which she is a part, and which she has reflectively endorsed. On her view, people who choose to engage in relationships that subordinate them to the will of others or that constrain their future choices will have their autonomy undermined (Oshana 2006: 99).<sup>46</sup>

For Oshana, an autonomous person must meet three general criteria: she must have certain cognitive capacities which make her capable of self-government (2006: 76-78), she must have *de facto* control over her own actions and choices (2006: 2-3), and these actions and choices must fall within a framework of rules (beliefs, values, principles) that she sets for herself

---

<sup>45</sup> Oshana puts this idea another way when she argues: "Autonomy requires that equilibrium of power be effected by the agent between herself and society" (Oshana 2005: 92).

<sup>46</sup> Oshana's account is global, unlike the other theories I've considered so far. This means that she is considering the autonomy of an agent over a long period of time, and not just the autonomy of a specific desire or decision. It is for this reason that she must conclude that any act which eliminates the possibility of future autonomy will render the agent non-autonomous.

(2006: 83-87).<sup>47</sup> This latter criterion requires that the beliefs, values, or principles by which a person directs her own life not be substantively incompatible with autonomy; a person cannot, on Oshana's view, autonomously choose to relinquish control of her own life.<sup>48</sup> In other words, a person must have the psychological capacity for autonomy, she must have actual control over her own life, and she must refrain from making choices incompatible with autonomy. These criteria are necessary and jointly sufficient, so if someone fails to meet them she lacks autonomy.<sup>49</sup>

For Oshana, there is an objective list of values or relationships which allows us to determine whether or not an individual is autonomous; some relationships are autonomy-makers and some relationships are autonomy-breakers. The Taliban Woman, for example, cannot be autonomous because her decision to participate in a religion that makes her servile to other persons is an autonomy-breaker (Oshana 2006: 120)<sup>50</sup>. Similarly, the subservient spouse is not autonomous because her decision to serve the needs of her husband before her own is an autonomy-breaker - even if she feels that this decision was autonomous.

This conclusion places Oshana's theory on the far end of the spectrum away from procedural, or content-neutral, theories like Friedman's and from weakly substantive accounts like Meyers'. At this point in the debate, positions were fairly polarized. One could adopt a

---

<sup>47</sup> Note the small role played by the authenticity condition. For Oshana's discussion of this point, see Oshana 2007.

<sup>48</sup> She argues: "But when people do not choose self-preservation, as in the case of voluntary slavery, restrictions on autonomously executed acts that eradicate one's dispositional autonomy can be upheld under a policy of judicious strong paternalism that even a staunch champion of individuality such as Mill would allow" (Oshana 2006: 117). But these restrictions do not just apply to voluntary slavery - they also apply to her four case studies: Taliban Woman, the happy slave, the browbeaten military recruit, and the subservient spouse (Oshana 2006: 120).

<sup>49</sup> Oshana appears to think that failure to meet any one of these criteria is enough to deem an agent non-autonomous.

<sup>50</sup> Another example Oshana uses is that of a person who is racially profiled. Someone who is a victim of racial profiling has what she calls frustrated autonomy, even if she subjectively identifies with racialized norms (Oshana 2005: 93).

strict, content-neutral approach like Friedman or one could adopt a substantive approach like Stoljar or Oshana. As the debate continued to mature, it looked like positions were becoming more and more entrenched, and the possibility of a *unified* feminist theory of autonomy was looking less and less likely.<sup>51</sup>

### **3. Breaking Dichotomies**

Part of the problem with debate entrenchment is that each opposing side continues to dig deeper and deeper into their own position; a chasm between views prevents unification, yes, but it can also prevent discussion. A further problem with such a harsh divide is that, eventually, both extremes start to look implausible. “Procedural or substantive?” like other sharp dichotomies (Deontologist or consequentialist? Liberal or illiberal? Realist or anti-realist?) fails to capture the wealth of possible views in the middle (perhaps one wants to be a quasi-realist, for example). Similarly, placing both poles on a simple spectrum fails to capture the many different axes on which a debate can rotate. (Where does virtue theory fall on the spectrum between deontology and consequentialism?) So far we’ve seen that, as the relational autonomy debate has matured, substantive theories have replaced procedural theories as the dominant view. Substantive theories are those that give significant weight to social factors in constituting autonomy, while procedural theorists argue that social factors contribute to (but do not constitute) what it means to be autonomous. Framed in this way, it is easy to see how the debate has polarized. One can either think that social factors are constitutive of autonomy, or not. The first view falls on one

---

<sup>51</sup> Paul Benson makes this point: “One of the interesting developments in the literature on relational autonomy over the past decade has been the emergence of what appear to be increasingly sharp and unbridgeable boundaries among different conceptions of autonomy and their respective evaluative commitments [...] The grounds for disagreement among competing accounts seem to be hardening” (Benson 2014: 88).

end of the spectrum, the latter is positioned on the polar opposite end, and never the twain shall meet. In this section, I comment on the way the debate has most recently changed, culminating in Veltman and Piper's 2014 volume, *Autonomy, Oppression, and Gender* (hereafter AOG). I suggest that dichotomies which have come to define the field – specifically the harsh divides between substantive and procedural theories and between feminist and mainstream theoretical approaches – are beginning to break down.

### ***3.1. Autonomy as Dialogical***

One clear way that the debate has begun to shift is that the alignment of constitutively social theories with substantive theories is breaking down. The first person to suggest this divorce was Andrea Westlund in her 2003 paper, “Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?”. She spells out her theory in more detail in her 2009 piece, “Rethinking Relational Autonomy”.<sup>52</sup> In this paper, she explicitly argues that autonomy can be constitutively relational without giving up its content-neutrality. Procedural accounts, on Westlund's view, can be constitutively relational if we understand autonomy as the disposition to hold oneself answerable to critical feedback. Autonomy is thus irreducibly social, since it relies centrally on dialogue with others, but it does not place any direct restrictions on the content of agents' desires, and so is not substantive (Westlund 2009: 28). Constructing her theory in this way allows Westlund to incorporate the strengths of both procedural accounts and substantive accounts. Like procedural accounts, Westlund can allow that women who endorse traditional values (like subservience) may well do so autonomously; she can avoid the often patronizing conclusions of substantive theories. But like substantive accounts, she builds a social element

---

<sup>52</sup> See also Westlund: 2012, 2013, 2014

into the definition of autonomy, itself, thereby recognizing the crucial role that our social relationships play in the formation and sustaining of our autonomy.

In his chapter in AOG, “Feminist Commitments and Relational Autonomy”, Paul Benson defends a dialogical theory as well, but his account is value-laden unlike Westlund’s procedural approach (Benson 2014: 89). Autonomy for Benson consists in an agent’s attitude of “*authority to speak or answer to others* for her choices and actions” (Benson 2014: 108). But he also places objective constraints on these attitudes: an agent’s attitudes toward her own status as a dialogue-participant must be formed rationally, and they must have sufficient evidence that they do, in fact, hold such a status (Benson 2014: 109). Agents must also be able to access socially available information regarding these attitudes, the attitudes can’t have been formed in a way that bypasses their rationality (presumably this condition rules out things like hypnosis), and the norms surrounding the agents’ attitudes must be, in principle, publicly shareable (Benson 2014: 109; cf. Benson 2005: 117-8). These four conditions put substantive restrictions on the types of attitudes of authority that appropriately constitute autonomy. Benson believes that in combining the answerability condition of dialogical approaches (like Westlund’s) with the restrictions of substantive approaches (like Stoljar’s) will allow him to take account of both agential *voice* and agential *authority* (Benson 2014: 110); strictly procedural or strictly substantive approaches, according to Benson, can only take account of one or the other.

### ***3.2. Utilizing Value***

Diana Meyers takes inspiration from Benson’s weak substantive answerability account. She argues that Benson’s account actually highlights that there are *two* axes on which to map how a theory of autonomy incorporates value. She refers to this thesis – that the two axes are

conceptually distinct – as the Double Axis Thesis (Meyers 2014: 115). The first axis, called the Directivity Axis, maps the extent to which a theory of autonomy utilizes value. At one end of the spectrum are value neutral accounts like Friedman’s which place no direct normative constraint on the content of an agent’s autonomous choices. At the far opposite end of the spectrum are value saturated accounts like Stoljar’s, Superson’s, and Susan Babbitt’s,<sup>53</sup> and in between these two poles are accounts which are “more or less value laden” (Meyers 2014: 116). Weakly substantive views would fall somewhere in between the two extremes, since they are value *laden* without yet being value *saturated*. The second axis is the Constitutivity Axis, which is composed of different constitutive value sets, which, according to Meyers, are “collocations of values that different philosophical accounts of autonomy explicitly or implicitly invoke to explicate the reflective procedure or motivational structure that renders choices and actions autonomous” (Meyers 2014: 120).

The basic idea behind splitting these two axes is that nearly all theories of autonomy are going to incorporate value somehow, at some stage of the game. Indeed, Meyers points out that even rational choice theory (“a paradigm of value neutrality”) makes use of epistemic values like consistency (Meyers 2014: 120). But this looks like a very different kind of value utilization than substantive theories which use value to dictate what sorts of desires can be autonomous (i.e. only those that foreclose oppressive values). Thus all theories of autonomy are going to fall somewhere along the Constitutivity Axis – even theories that are classified as value neutral on the Directivity Axis. Benson’s theory, which he classifies as weakly substantive, is actually more closely aligned on the Directivity Axis with value neutral theories, since “it does not affirm that the extent or the gravity of women’s compliance with oppressive gender norms is decisive

---

<sup>53</sup> See Babbitt 1993; Stoljar 2000, 2014; Superson 2005. Note that there is some ambiguity over how to classify Stoljar’s theory (Meyers 2014: 117-8; see my Chapter 4).

with respect to their autonomy, and it asserts that women can autonomously comply with such norms” (Meyers 2014: 119; cf. Benson 2005: 136). On the Constitutivity Axis, Benson’s theory utilizes a specific value set, including the psychological values of self-confidence and self-worth – both of which are needed for autonomous action, but neither of which place direct restraints on the content of a woman’s desires or choices (Meyers 2014: 119).

The advantage of the split, from a feminist perspective, is that feminist autonomy theorists can still uphold value neutrality (which Meyers argues we have good reason to do) without claiming (falsely) that the theories don’t appeal to values at all. Meyers’ Double Axis Thesis thus represents a clear break from the strict dichotomy between substantive (value-full) theories, on the one hand, and procedural (value-less) theories on the other.

### ***3.3. Mainstreaming Feminism***

I want to end this review on a positive note. When Mackenzie and Stoljar published *Relational Autonomy* in 2000, their overriding goal was to salvage the concept of autonomy for feminist purposes. The blossoming of the debate in the past fifteen years, which I have discussed in this chapter, demonstrates clearly that they achieved this goal. But as the feminist debate has developed, it has also influenced the mainstream debate; indeed, it is becoming more and more difficult to state, definitively, that an account is either feminist or mainstream. The impact of the social on both persons and autonomy has become a central consideration for autonomy theorists on both sides of the ‘divide’.<sup>54</sup>

I will describe one example of what we might call the *mainstreaming phenomenon*, in the debate between Steven Weimer and Michael Garnett regarding the social nature of selfhood and

---

<sup>54</sup> See, for example, the conversation between Steven Weimer and Michael Garnett in Garnett 2013, 2014, 2015 and Weimer 2014; Killmister 2013a, 2013b.

autonomy. In his provocatively titled 2013 paper, “Taking the Self out of Self-Rule,” Michael Garnett argues against what he calls “the deep self view of self-rule” according to which an agent is self-ruled if she is ruled by her deep, or authentic, self (Garnett 2013: 23). Instead, Garnett proposes that autonomy is better understood as a lack of rule by others agents, a view he calls “the negative social view of self-rule” (Garnett 2013: 27). In arguing for this position, he explicitly acknowledges that he is breaking with “traditional” ways of thinking about autonomy *even more radically* than relational theories of autonomy like Meyers’ (Garnett 2013: 27).<sup>55</sup> His account, unlike Meyers’ is constitutively social.<sup>56</sup> In fact, Garnett argues, “On the negative social view presented here, one’s autonomy is *purely* a matter of the type of social control to which one is subject and *need have nothing especially to do with the quality of one’s agency*” (Garnett 2013: 27, emphasis added). He expands his account in his 2014 paper “The Autonomous Life: A Pure Social View.” Here he argues that conditions we normally associate with autonomy (like reflective endorsement, for example) are only autonomy traits – they are not necessary or sufficient for autonomy (Garnett 2014:147). Instead, an agent only needs enough of these traits to ward off rule by others. This *certainly* distinguishes his account from Meyers’!<sup>57</sup>

---

<sup>55</sup> Because the line is blurring so much, I think it’s not inappropriate to say “traditional” instead of traditional.

<sup>56</sup> I should note that Garnett identifies Friedman as a theorist who relies on a constitutively social account of autonomy and a constitutively social account of the self. He then argues that his own theory is an improvement upon her theory, since he does not have to rely (tacitly) on the deep self view. Friedman, on the other hand, thinks that a social account of the self will generate a social account of autonomy, since autonomy is really rule by the authentic self. Now, he is correct that Friedman relies on authenticity conditions for her theory of autonomy, but her account is *not* constitutively social. Nor is it clear that she relies, any more than other feminist theorists, on a constitutively social account of the self. I would attribute this combination of views to someone like Marina Oshana, who also dismisses accounts that rely too heavily on authenticity (2007). In other words, although the line between traditional and feminist theories is blurring here, I’m not sure the feminists are being accurately represented.

<sup>57</sup> Although the identification of autonomy traits does sound similar to Meyers’ discussion of autonomy competency.



One of the problems with Garnett's account, pointed out by Weimer, is that the list of autonomy traits includes several conditions that are actually external to the agent. Garnett's list includes the following traits: self-creation, self-sufficiency, intelligence, a sense of self-worth, and a capacity for critical reflection. All of these traits are arguably internal to the agent. But his expanded list includes rights of democratic participation, noncoercion, basic levels of material goods required for survival, and the love of family and friends (Weimer 2014: 160). This latter list is composed of things that are agent-external. So, Weimer asks, if all that is required for autonomy is a sufficient set of autonomy traits, does this mean that an agent could possess only external traits? If not, Weimer argues, "the account falls prey to an adjusted version of an objection Arneson leveled [...] resistance to foreign rule does not imply autonomy because it may be that 'no one is ruling at all ... a person not [easily] steered by any external agency might be just rudderless, rather than autonomous'" (Arneson 1994: 55; Weimer 2014: 161). Such an account might lead to wildly unintuitive results. To avoid this objection, Garnett must require that, of the list of sufficient traits, at least one must be internal to the agent. Weimer argues, however, that the only reasons we think such internal traits are necessary depend upon a tacit acceptance of the very same deep self view that Garnett rejects (Weimer 2014: 163). Thus, Garnett must either accept wildly unintuitive results, or accept that he is tacitly committed to the deep self view of autonomy, after all.

This debate and Garnett's 2015 response highlight perfectly the influence of feminist accounts of autonomy on the mainstream debate. Questions that motivated theorists like Natalie Stoljar to propose and defend her feminist intuition are now motivating mainstream theorists. This is the final dichotomy that I see breaking down in the recent relational autonomy debate. Feminists have not only, then, revived the notion of autonomy for their own purposes, but have

in doing so brought attention to serious problems like internalized oppression which are now being taken up in the mainstream debate. As the debate moves forward, I hope this interaction continues.

## CHAPTER IV

### NEITHER ISLANDS NOR DUPES: RELATIONAL AUTONOMY AND THE SELF

The conflict between traditional and feminist conceptions of autonomy is not a new one. The arguments against traditional accounts often focus on the charge that these accounts are individualistic -- that they celebrate an individual separated from familial or social ties. Feminists have been quick to point out that such an atomistic model of the individual does not represent the real lives of women (or men, for that matter). Any theory of autonomy that does not take into account social relations is seriously flawed. In order to rectify these weaknesses, feminists have presented an alternative set of theories which take into account our social embeddedness. Autonomy, on these accounts, is relational. It is not an ideal of separateness, but instead a concept that incorporates the many different relationships of which selves are a part.

To be sure, relational accounts are a welcome change from the model of "I am a rock, I am an island." But I wonder if we haven't left something behind. I wonder if, in our hurry to correct the model of autonomy which told us to "build walls, to build a fortress deep and mighty that none may penetrate..." we have abandoned the "self" in self-government. In this chapter, I challenge feminist autonomy theories which conceptualize the self and autonomy as socially constituted.<sup>58</sup> These theories are explicitly motivated by the normative commitments of feminism, but I argue that a conception of selves and autonomy as constitutively social is ill suited to meet feminist goals. I end by suggesting that a more individualist theory – of both

---

<sup>58</sup> It is no longer possible to talk about "feminist theories of autonomy" as if this were a homogenous group. Feminist theorists now occupy the entire range of positions (from purely procedural to strongly substantive) evident in the mainstream literature. For this reason, I am narrowing my target to just those feminist positions that are substantive (value-laden instead of value-neutral). I engage with a variety of feminist views when providing the background of the debate, but all other references to "feminist theorists of autonomy" should be understood as references to feminists with substantive theories.

selves and autonomy – better aligns with feminist commitments to ending the oppression and subordination of women.

### **1. Why Relational?**

The feminist defense of relational autonomy takes its point of departure from these claims: first, that traditional theories of autonomy presuppose an individualistic conception of the self; and second, that these theories posit an individualistic conception of autonomy. The problem with both types of individualism, it is argued, is that they prevent an accurate understanding of the social realities of persons. These theories begin with feminist social/political aims (such as the eradication of women’s oppression) and then frame a theory of autonomy to suit these aims. This suggests that what matters for a feminist theory of autonomy is strictly its fit with feminist social/political aims (where fit is understood as some sort of positive contribution to their realization). On these accounts, the right theory of autonomy is the one that best fits with feminist social/political aims, and the theory that best fits is the one that most effectively contributes to the realization of feminist goals – specifically, the eradication of women’s oppression.

### **2. Socially Constituted Selves**

However, it is simply not the case that all traditional theories of autonomy are built on a foundation of robust individualism. Some early feminist critiques charged that traditional accounts of autonomy relied upon a conception of the self as a *self-maker*. Lorraine Code, for instance, argues against a hyperbolized version of autonomy, which she argues “descriptively configures and prescriptively animates the discourses of self-sufficient individualism in which

‘autonomous man’ retains his place as an iconic figure, emblematic of an unrealistic imperative toward self-reliant self-making” (Code 2000: 183).

But Joel Feinberg and Gerald Dworkin, for example, argue that the concept of a self-made man or woman is incoherent. To suppose that autonomy requires self-creation *ex nihilo* is to dictate that *no agent can ever be autonomous*.<sup>59</sup> We are not isolated and self-creating; we are, as Feinberg notes, *social* (Feinberg 1989: 45). Any conception of the self as (what Dworkin calls) "the unchosen chooser" or "the uninfluenced influencer" is mistaken (Dworkin 1989: 58). Even theories like Harry Frankfurt's, which don't acknowledge a social self, only fail to do so contingently; that is, there is nothing about traditional theories of autonomy that prevent thinking about the self as at least somewhat relational (Frankfurt 1971; Friedman 2003; Westlund 2009: 26).

But this may not be a sufficient response. As Jennifer Nedelsky rightly points out, a theory must do more than acknowledge that "individuals will, of course, encounter one another" (Nedelsky 1989: 9). It is not enough to say that selves exist in societies. We might even think that to say this is to miss the point, since it presupposes a sharp distinction between individuals and the societies in which the individuals exist - and this distinction, or boundary, is exactly what is at issue (*ibid*).

### ***2.1. Social Causation***

So what exactly are the desiderata a conception of the self must meet in order to be sufficiently social? I propose that there are at least three candidates for understanding feminist views on this

---

<sup>59</sup> Joel Feinberg, for example, argues that to require self-creation is to "conceive of authenticity in such an exalted way that its criteria can never be satisfied, or else to promote the ideal of authenticity in a self-defeating way."

topic, which correspond to three ways to interpret what we might call the *Social Self Thesis (SST)*:

**SST(1).** Individuals are social in the sense that they always develop in social environments; this is a *causal* claim.

**SST(2).** Individuals are social in the sense that they are at least partially constituted, in an ongoing way, by their social relationships; this is a claim about the *social constitution* of the self.<sup>60</sup>

**SST(3).** Individuals are social in the sense that they are always determined by their social relationships; this is a claim about the *determination* of individual identity by social factors.

Much of the original feminist criticism expressed **SST(1)**. Feminists argued that human development always requires relationships of dependency with others; Annette Baier, for instance, has argued that persons are always second persons, that is, persons who “grow up with other persons,” who are “long enough dependent upon other persons to acquire the essential arts of personhood” (Baier 1981: 180). But **SST(1)** claims, at most, that people are socially *caused*; it does not claim that people are socially *constituted*. To say that a person is socially caused is only to say that there was a time during which she was dependent on others. But this is trivially true.<sup>61</sup> In other words, if the feminist criticism of traditional autonomy theories is just that they don’t recognize **SST(1)**, this request could easily be accommodated without altering the

---

<sup>60</sup> This is how Marilyn Friedman interprets Jennifer Nedelsky’s claim (for example) that selves are socially constituted. See Friedman 2014: 56.

<sup>61</sup> At least at this point in the debate, this fact should be taken as trivially true. I am not discounting the work that care ethicists have done to draw the fact of dependency to our attention.

traditional theories at all (Mackenzie and Stoljar 2000: 7). This gives us reason to conclude that **SST(1)** is not what feminist autonomy theorists mean when they say that the self is social.

This conclusion is especially warranted since the feminist criticism is not simply that traditional autonomy theorists don't recognize that selves are socially *caused*. Even if traditional theorists were to accept **SST(1)** explicitly, one might think that **SST(1)** is still too weak to ground a feminist theory of autonomy. This is at least in part because it is compatible with a theory of autonomy that requires agents to transcend social influence and become self-sufficient.<sup>62</sup> Such accounts could acknowledge that we all are at one time dependent, while still requiring that we somehow rise above this dependency. It is true that part of the feminist critique has been to highlight that such a requirement fails *descriptively*. Linda Barclay, for example, points out that dependency is not something we transcend, but rather something that is ongoing. "We do not merely acquire autonomy competency in childhood and then become fully independent," she argues (Barclay 2000: 58).

But what I want to consider is the *normative* failures of **SST(1)**, not the descriptive failures. I want to suggest that the feminist criticism of **SST(1)** is based on the idea that, because it is consistent with the requirement that we somehow transcend social influence, it is not attendant to the pervasive nature of women's oppression. It is worth quoting Catriona Mackenzie at length here, since I think she provides a clear explication of this point:

If persons are socially constituted, then external conditions, including our social relations with others, shape the process of practical identity formation – the *self* of self-governance [...] On the one hand, this is not necessarily problematic from a relational perspective [...] On the other hand, in contexts of social oppression, as is shown by the phenomenon of adaptive preference formation, severely limited opportunity conditions can *deform the process of practical identity formation* and impair the development and exercise of autonomy competence (Mackenzie 2014: 31; emphasis added).

---

<sup>62</sup> For this criticism of 'individualist' conceptions of autonomy see Mackenzie 2008: 520.

The point here is that oppressive relationships impact not only an agent's autonomy, but also the very construction of the agent's identity itself (see also Mackenzie 2008: 513). **SST(1)**, then, which allows for the possibility that selves can transcend social influence, fails to take into account that the 'selves' in question are themselves molded by these social influences, so that they cannot be transcended. This failure, in turn, is a failure to understand that oppression is not just something that *happens* to an agent, but rather something that helps *form* her. In order to conceptualize oppression in a way that takes seriously the formative role that it plays, feminists have dismissed **SST(1)** in favor of a more robustly relational account of the self.<sup>63</sup>

## ***2.2. Social Constitution***

But what hasn't been made clear in the debate so far is to what *extent* selves are social. Taking a closer look at the specific language in this debate suggests that **SST(2)** is a better representation of what it means for a self to be social. Perhaps selves are social in the sense that they are partially constituted by their social relationships. Such an interpretation could admit of many degrees. Perhaps it has something to do with an agent's own self-conception. Jennifer Nedelsky, at times, suggests that the debates over what it means for a self to be social are at bottom debates about how an agent thinks of herself (Nedelsky 2011: 56). She also argues that our identities are not comprehensible to us in isolation from our relationships (Nedelsky 2011: 121). Paul Benson argues that our social relationships mold "who we regard ourselves as being"

---

<sup>63</sup> Diana Meyers also explains this position well: "To be unjustly subordinated, it would seem, is to be diminished in one's selfhood *and* to have one's agency curtailed. Otherwise, what's the harm?" (Meyers 2010: 9, emphasis added).



(Benson 2000: 85). And John Christman argues that we only experience ourselves in relation to certain others or as occupying certain interpersonal roles (Christman 2009).<sup>64</sup>

But some of the language utilized in this debate suggests that this may not be robust enough to capture what it means for a self to be social. Part of the problem with understanding the *Social Self Thesis* as a claim about the *influence* of social factors, it seems, is that it relegates social influences to the level of contingency. But some relational theorists question this contingency, and argue instead that social factors are *necessary components* of human beings. SST(2) could also express the view that social influences are *necessary* components of persons, so that people are irreducibly social not only in their self- conceptions, but in their very nature. On such an account, it would be impossible to conceive of a person abstracted from social influences. Jennifer Nedelsky, for example, argues: “We come into being in a social context that is *literally constitutive* of us” (Nedelsky 2011: 120). She also argues: "Relations (at all levels) are then not just the 'conditions' under which a freestanding self emerges. It is the very nature of human selves to be in interaction with others. In important ways, *they do not exist apart from these relations*” (Nedelsky 2011: 55).<sup>65</sup>

Indeed, when we remember that relational autonomy was originally proposed as an alternative to the image of an atomistic, hyper-individualistic self-creator, it seems clear that SST(2) must express something stronger than just social influence. Natalie Stoljar seems to agree, when she points out that relational autonomy theorists have adopted themes developed in the work of Charles Taylor, who defends a communitarian conception of the self according to which “we define our identity always in dialogue with, sometimes in struggle against, the things

---

<sup>64</sup> See also: Mackenzie and Stoljar 2000: 7-8.

<sup>65</sup> She argues further that her relational approach “disrupts” the distinction between self and others (2011: 376).

our significant others want to see in us” (Taylor 1994: 28; Stoljar 2011: 378). On Taylor’s interpretation, the self is dialogical in its very nature. Taylor argues against what he calls the “monological ideal,” according to which “we need relationships to fulfill, but not to define, ourselves” (Taylor 1994: 33). This ideal, according to Taylor, “seriously underestimates the place of the dialogical in human life” (ibid); our identities “crucially depend” on our relationships with others (Taylor 1994: 34). Taylor’s own position corresponds to **SST(2)**, according to which social factors play a role in constituting our identities. On Taylor’s view, and on Nedelsky’s, the social plays an ongoing role in constituting our identities. In other words, the social has a continuing influence. This may be the most appropriate interpretation of what it means for the self to be socially constituted.

### ***2.3. Social Determination***

**SST(3)**, on the other hand, looks too robust. I’m not sure there are any feminist theorists of autonomy who explicitly rely on a view of the self as wholly determined by social factors. Indeed, what would *autonomy*, or self-government, be on such an account? It looks like there must be a middle ground, which I’ve tried to isolate as **SST(2)**, an account of social selfhood which is a stronger version of **SST(2)** but which does not yet grant determinative power to social forces. Something like **SST(2)**, which grants social forces a stronger influencing power than **SST(1)** looks necessary to capture the realities of human identity for some feminist autonomy theorists; and it is also, on these views, the necessary foundation for any theory of autonomy that hopes to capture the pervasive nature of women’s oppression.

It will help to provide an example to illustrate the motivation behind this discussion, and to show why I think it is necessary to explicate exactly what *social constitution* means. In “Who

Are We?: Modern Identities Between Taylor and Foucault,” Allison Weir presents a complicated conception of self-identity according to which we are socially and historically constituted. Our identities, she argues, “*are* our connections” (Weir 2009: 551, emphasis added). When we think of ourselves as self-creating individuals, or when we separate the ‘I’ from the ‘We’, we bind ourselves to an unchosen identity. The only way we can be said to choose who we are is if we recognize the extent to which our identities are social. She argues that “being truly oneself does not mean accessing some pre-given core of self; it means participating fully in one’s connections” (Weir 2009: 543). It is worth quoting her at some length:

I am making the strong claim that to the extent that my important social and political identities are not salient to me, to the extent that I fail adequately to examine what these mean, to that extent I will *fail* to recognize the *truth* of myself and my relations to others, and I will *fail* to progress toward freedom from fixed identities. This means that if we think our identities are just about what is salient for me, what I *choose* to identify with, then we are failing to understand who we really are. The fact that we are constituted through social contexts and through regimes of power means that we do not know ourselves until we understand those contexts and those power regimes (Weir 2009: 542-3).

I would argue that Weir endorses a strong version of **SST(2)**. She does not yet want to grant that our identities are wholly determined by our social relations, but she verges on endorsing such a view. I worry that, *if this is what it means for selves to be social*, then our social relations have far too much power over our identities, and this will have dramatic results for how we understand autonomy. In the next section, I go on to discuss specific worries with the social constitution of selves.

### 3. Concerns with the Social Constitution of Selves

#### 3.1. Questioning Social Constitution

I propose that we have reason to question the strong reading of **SST(2)**, according to which social relations play an ongoing role in constituting our identities (and not just our self-conceptions). The first, of special concern to feminists, is that defining people by their social relations can itself be oppressive.<sup>66</sup> Historically women have been defined as mothers, wives, caregivers, homemakers; in short, they have been defined by their relationships. It is precisely because the social context of so many women is oppressive that defining women in terms of their social context risks defining them in terms of their own oppression - and this is something that feminists have tried to overcome.<sup>67</sup> Women are *not* just mothers or teachers or caretakers; they are individuals with unique interests and concerns.<sup>68</sup> Moreover, blurring the line between the self and others can be dangerous.

The second concern with **SST(2)** is that since our social relationships are often given to us and not chosen, and since (according to this robustly social conception of selves) these relationships constitute our identity, it looks like *we do not choose who we are*.<sup>69</sup> In other words, if human beings do not exist apart from their relationships (like Nedelsky suggests), it is unclear how we would have any ground from which to critically examine - or change - these relationships. Such critical examination is required if we are to address internalized oppression.

---

<sup>66</sup> Feminists like Ann Cudd have pointed out that: "Social forces are the very forces of unfreedom in many cases" (Cudd 2012: 379).

<sup>67</sup> Nedelsky 2011: 122.

<sup>68</sup> Linda Barclay, for example, points out: "Traditionally, women's lives have been devoted to the care of others, and if anything the problem has been to find a space for the expression and pursuit of one's own interests" (Barclay 2000: 59).

<sup>69</sup> Here is where my point from the introduction reappears, regarding the loss of the self in self-government.

What I am questioning, in short, is the idea that *for the purposes of constructing a theory of autonomy*, we should think of selves as irreducibly social; that a person's very identity, even if not wholly determined by her social context, is still overly influenced by it such that if this context were to change, *she would not be the same person*. Social context, of course, plays an immense role in the formation of our self-conceptions. But I question whether **SST(2)** is the best way for feminists to define the self in this debate. Accounts that give social factors too much power over the construction of our identities may forget that it is often these factors that oppress us in the first place.

There is a quick response to the worries I have just expressed, namely, that I have merely pointed out unpleasant normative implications of accepting **SST(2)**. But, one might think that normative implications of any theory are irrelevant if the theory relies on a descriptive falsehood. Wouldn't it be great if we were not, in fact, determined by our social relationships? Wouldn't that make combatting oppression much easier? In fact, there are many things that would make combatting oppression much easier; but if these things exist only in fiction, then why speculate?

This is a complicated question, and ultimately I'm not convinced that autonomy theorists have to provide a final answer to the question of *how far down* the social influence on selfhood goes; it is certainly not my intention here to do so. I want only to question the idea that it is better for autonomy theorists to think of selves as robustly social in this way. I want to question the idea that any theory of autonomy that hopes to vindicate the normative commitments of feminism must rely on a conception of selves as irreducibly social.

### 3.2. *What Are Social Relationships?*

To clarify my point, it is necessary to pin down exactly what we mean by ‘social relationships’ in this debate. When feminist autonomy theorists talk about the social constitution of selves, are we talking about being social in the abstract (which may in fact be a necessary feature of persons) or about being social in regard to particular social relationships with particular people? The problem with **SST(2)** is that it is unsatisfactory on both accounts. If being social means that we are defined by particular relationships with particular others, then it becomes difficult to account for the ways people change over time. This is what John Christman has in mind when he argues that conceiving of selves as constituted by their social relations, “runs the risk of ignoring the very variability, contingency, and temporally fluid nature of human existence that motivates the rejection of old-style individualism” (Christman 2004: 145). Consider, for example, a woman who is very close with her mother. There is a sense in which this relationship constitutes her identity. But it cannot wholly constitute who she is. If this were the case, the woman in question would cease to exist when her mother passed away. And although she might be deeply affected, she is still *herself*. This cannot be the sense in which selves are said to be socially constituted. But there are also problems if we understand **SST(2)** in the abstract. Now a person is not defined in relation to particular others, but rather in relation to a broader social setting. But on such an account, women’s identities (so long as patriarchy survives) are constituted by oppression; this is problematic for reasons I will go on to discuss.

I want to suggest that **SST(1)**, which acknowledges the influence of social factors on selfhood but does not go so far as to say that these factors are determinative of selfhood, is in fact relational enough to live up to the normative commitments of feminism. This view occupies a middle ground between traditional theories of autonomy, which do not emphasize the role that

social factors have in self-formation, and robustly relational accounts of autonomy, which go too far in the opposite direction.

#### **4. Socially Constituted Autonomy**

In the previous sections, I focused mainly on clarifying the terms of the feminist autonomy debate. What do we mean when we say that selves are social? Is it necessary for feminist autonomy theorists to defend a conception of selves as irreducibly social? I argued that it might be better, for the purposes of this debate, to acknowledge the influence of social factors on selfhood, without granting such factors so much power over our identities. For the remainder of this chapter, I shift focus from the social constitution of selves to the social constitution of autonomy.

First, I present the different ways in which autonomy itself could be social. Next, I present objections against accounts which allow social factors to play a decisive role in determining an agent's autonomy status. I then take a closer look at feminist strong substantive autonomy theories to illustrate the way a strong social account functions. Substantive accounts – contrasted with content-neutral accounts – are those which place normative constraints on some aspect of an agent's autonomy. Weak substantive accounts may include weak normative substance – for instance, such accounts might include self-respect as a required condition for autonomy. *Strong* substantive accounts, on the other hand, place *direct* normative constraints on autonomous choices or desires (Benson 2005, 133; cf. Stoljar 2014b, 34). It is with this latter type of theory that I take issue. Ultimately I suggest that my objections provide reasons to question strong substantive accounts. And, since the specific strong substantive accounts I'm considering in this chapter – those of Natalie Stoljar, Sonya Charles, and Anita Superson – are

also robustly relational, I also argue that we have reason to question robustly relational accounts as well.

#### ***4.1. Social Autonomy: Two Interpretations***

There are two ways to understand what we might call the **Social Autonomy Thesis (SAT)**.

**SAT(1).** Social factors play *a role* in determining the autonomy status of agents; such factors are *causally* related to an agent's autonomy.

**SAT(2).** Social factors play *a defining role* in determining the autonomy status of agents; such factors are *constitutively* related to an agent's autonomy.

The difference between **SAT(1)** and **SAT(2)** is that the latter includes social factors in the defining features of autonomy itself. The influential accounts of Gerald Dworkin and Joel Feinberg, for example, exemplify **SAT(1)**. A Dworkonian or Feinbergian theory could readily acknowledge that social factors *influence* the ability of an agent to examine her own desires, for example, but such factors would not form part of the definition of what it means to be autonomous.

Now consider feminist theorists like Natalie Stoljar and Anita Superson, who argue that the right kind of social factors are necessary in order for an agent to be autonomous. They reason that a patriarchal society can lead women to adopt values that undermine their own autonomy – even if such values are reflectively endorsed by the agent herself. Oppression is not just an external force, but a force that becomes internalized by the oppressed agents so that they come to accept oppressive preferences as their own (I am thinking here of cases like Thomas



Hill's Deferential Wife).<sup>70</sup> The problem is that according to traditional theories like Frankfurt's, or theories that adopt **SAT(1)**, these *oppressive preferences* seem like *legitimate preferences*.

The strategy of feminist strong substantive theorists has therefore been to argue that actions deriving from (or influenced by) oppressive preferences are not, in fact, autonomous.<sup>71</sup> But this becomes difficult to argue on traditional individualist accounts according to which autonomy just consists of an individual agent rationally reflecting on her preferences. Oppressive preferences can survive this *individual process*.<sup>72</sup> But if autonomy is, instead, constituted by an individual's social relations, it becomes easier to argue that her oppressive preferences are non-autonomous. Why? Because she exists in the *wrong sorts* of relationships. Natalie Stoljar argues that "preferences are autonomous if and only if they correspond to moral or otherwise objectively correct features of the world" (Stoljar 2014b: 35). The wrong sorts of relationships are thus those which do not correspond to objectively correct relationships, which on these accounts would presumably be those in which one partner is not unflinchingly subservient. Autonomy is thus a trait of persons which consists in the presence of certain

---

<sup>70</sup> Hill, Thomas E. Jr. "Servility and Self-Respect," in *Dignity, Character, and Self-Respect*, ed. Robin S. Dillon. New York: Routledge, 1991 (reprinted from *The Monist* 57 (1973)) pp. 76-92. Thomas Hill describes the Deferential Wife as a wife who: "Buys the clothes [her husband] prefers, invites the guests he wants to entertain, and makes love whenever he is in the mood... She does not simply defer to her husband in certain spheres as a trade-off for his deference in other spheres. On the contrary, she tends not to form her own interests, values, and ideals, and when she does, she counts them as less important than her husband's" (Hill 1991: 5).

<sup>71</sup> I am not considering Oshana's social-relational account in this chapter as an example of a strong substantive account. Oshana is concerned with global autonomy, not local autonomy, so she thinks it could be acceptable for a person to choose – autonomously – to be a submissive wife; but this person has (by definition) deprived herself of the possibility of future autonomy. She argues: "Even if a person can autonomously select a life that denies him self-determination, the result of that choice will be a situation in which the chooser lacks autonomy" (2006, 56). In this respect, Oshana's account does not differ from a proceduralist account. It does make her position different from the other feminist substantive accounts I'm considering here (like Stoljar's, Superson's, and Charles'), even though she is often aligned with feminist substantive theorists.

<sup>72</sup> Indeed, this is why Stoljar draws our attention to the Feminist Intuition.

appropriate social relations (like egalitarian marriages) and the absence of those that would undermine autonomy (like the marriage between the Deferential Wife and her husband). Relationships that deny an agent control over her social circumstances are by definition, on accounts that hold **SAT(2)**, incompatible with autonomy. For reasons that I explain below, I question whether or not **SAT(2)** is the best way to construct a feminist theory of autonomy.

## **5. Concerns with the Social Constitution of Autonomy**

In this section, I present arguments against the idea that autonomy is exhaustively socially constructed. I should note that the burden of proof is on me. Because my arguments in this chapter can ultimately support a theory of autonomy that is more individualistic than social, I have to demonstrate that my account is an acceptable feminist account of autonomy, even though it is less robustly relational. Can this be done? One way to proceed is to note that there is a potential tension in constitutively relational accounts of autonomy. This tension highlights an ambivalence regarding the level of individualism that feminists find acceptable in theories of autonomy. It turns out that even constitutively, or robustly, relational theorists still want to maintain a sort of individualism; but this is not possible unless autonomy is less relational than is specified by such accounts.

### ***5.1. Deformed Desires***

Part of the problem with internalized oppression is that it produces adaptive preferences, or deformed desires. Such desires are deformed because they are (arguably) not in the best interest of the agents who hold them; they are, instead, in the interest of the oppressor. Anita Superson argues, for example, “Slavish values aim only at satisfying another, typically to the disadvantage

of the ‘slave’” (Superson 2005: 110). A woman who adopts deformed desires (or slavish values) thus adopts desires that serve to reinforce her own oppression, which is what leads the feminists considered here to conclude that these agents are mistaken about what is good for them – *objectively speaking* (Superson 2005: 121). But because these agents profess that the desires are really their own, feminist substantive theorists have formed the general principle that desires formed under a system of oppression, and that align with oppressive norms, are thereby non-autonomous; in other words, they discount the agent’s subjective perspective on her own autonomy. A woman who argues that she *really does* desire to be a Deferential Wife, for example, is confused about her own good because she has formed her desire under an oppressive system. Her desire for deference is thus non-autonomous.

But when we consider that women in a patriarchal society are (to greater or lesser degrees) oppressed, we get a strange sort of result. If “being formed under conditions of oppression” is the criterion for non-autonomy, and women are oppressed...does this mean that women’s desires cannot be autonomous? If this is the case, and if these desires align with oppressive norms, then what we have is a feminist account of autonomy that allows autonomy for the oppressor but not the oppressed. This result, in turn, means that we understand women in patriarchal societies as oppressed and therefore non-autonomous, regardless of their own subjective relationship to the desires they hold, and regardless of whether or not the specific desire adoption in question was actually motivated by their oppressive circumstances. In short, on this line of thought, as women living in a patriarchal society, we are bound to be dupes.<sup>73</sup>

---

<sup>73</sup> The language of “dupes of patriarchy” comes from Uma Narayan: “Minds of Their Own: Choices, Autonomy, Cultural Practices, and Other Women,” in *A Mind of One’s Own: Feminist Essays on Reason and Objectivity*, edited by Louise M. Antony and Charlotte E. Witt. Boulder, Colorado: Westview Press, 2002.

## **6. A Closer Look at Feminist Strong Substantive Autonomy**

But surely all this means is that we need a clearer explication of what it takes for a desire to be deformed; perhaps it is not the case that a particular desire's being "formed under conditions of oppression" is sufficient for the denial of autonomy. In this section, I take a closer look at three feminist strong substantive accounts of autonomy – those offered by Natalie Stoljar, Sonya Charles, and Anita Superson. After a brief presentation of each view, I extract a detailed account of deformed desires in order to demonstrate that, even with a clearer explication of what it takes for a desire to be deformed, strong substantive theories are susceptible to my objections. More specifically, I argue that their treatment of deformed desires shows what I take to be the problem with strong substantive accounts. Because they rely on constitutively relational accounts of autonomy, they yield a result which is normatively unsound from the very feminist perspective their theories were designed to reflect.

### ***6.1. Three Strong Substantive Accounts***

In "Autonomy and the Feminist Intuition," Natalie Stoljar argues that neither content-neutral nor weak substantive accounts of autonomy are able to uphold the feminist intuition, which claims that "preferences influenced by oppressive norms of femininity cannot be autonomous" (Stoljar 2000: 95). Stoljar utilizes Kristen Luker's study of women who engaged in contraceptive risk-taking as an illustration of internalized oppression (Stoljar 2000: 99). Autonomy theories which only examine factors internal to the agent cannot separate oppressive preferences from legitimate preferences, and so cannot conclude that Luker's women are non-autonomous – even though Stoljar claims that they clearly incite our feminist intuition. Only a strong substantive theory is

able to account for the heteronomy of Luker's women, because only a strong substantive theory examines the *content* of the norms they endorse.

Sonya Charles takes up and defends Stoljar's strong substantive account in her 2010 paper, "How Should Feminist Autonomy Theorists Respond to Internalized Oppression?" She argues, like Stoljar, that only a strong substantive account is capable of distinguishing between pernicious socialization on the one hand, and non-problematic, or acceptable, socialization on the other (Charles 2010: 413). She utilizes Thomas Hill's example of the Deferential Wife, and her own example of the Surrendered Wife as test cases, and finds that content-neutral theories must conclude that both wives are autonomous. The Surrendered Wife is based on a book by Laura Doyle of the same name, which defends a lifestyle where "the control women wield at work and with children must be left at the front door of any marriage to revitalize intimacy" (Charles 2010: 417). Charles thus provides the Surrendered Wife as a real life example of the Deferential Wife. She argues that content-neutral theories must conclude that both wives are autonomous, because oppressive and non-oppressive socialization are functionally similar in the procedure for autonomous decision-making (Charles 2010: 424). But Charles thinks she has provided adequate evidence that the women in these two cases are non-autonomous. They have both internalized problematic norms of femininity and are thus both non-autonomous. In her view, only a strong substantive theory is capable of reaching this correct conclusion.

Both Stoljar and Charles exemplify the methodology I discussed earlier, namely, they begin with an intuition about who does or does not count as autonomous, and then construct a theory to reflect this intuition. Both classify their own theories as strongly substantive. This classification suggests that it is the *content* of an agent's beliefs or desires that is most relevant for determining her autonomy status. But both Stoljar and Charles discuss a second factor,

namely the *impact* that such beliefs or desires have on the agent's perception of her own self-worth. Charles argues, for example, that her theory highlights how "decisions that reflect a certain devaluation of self (or lack of self-worth) should not count as autonomous" (Stoljar 2010: 413). Stoljar argues that, "Women who accept the norm that pregnancy and motherhood increase their worthiness accept something *false*" (Stoljar 2000: 109). So far this is consistent with a strong substantive theory. But she continues, "And because of the internalization of the norm, they do not have the capacity to perceive it as false" (Stoljar 2000: 109).

The proper classification of these theories isn't important here.<sup>74</sup> But it is important to determine *what is doing the work* on Stoljar and Charles' accounts. Is it the *content* of an agent's beliefs that determines her autonomy status, or is it the *impact* that such beliefs have on her normative competence? For both Stoljar and Charles, I think it must be the content doing the work. Even though, in more recent work, Stoljar suggests that the case of deformed desires may be "more complicated" than strong substantive theorists like herself have previously thought (Stoljar 2014a: 252), she still concludes that "many of the adaptive preferences of concern to feminists are autonomy impairments" (ibid).<sup>75</sup> Citing the impact of false beliefs on an agent's normative competence might provide additional justification for denying her autonomy, but this impact doesn't play the key role in Stoljar or Charles' accounts. The content of an agent's desires or beliefs – whether or not she is committed to her own subservience – is what determines whether she is autonomous.

---

<sup>74</sup> I'm not sure if Stoljar's most recent work suggests that she wants her own account to be a normative competency account rather than a strong substantive account. As I say, I don't think the classification ultimately matters for my purposes here.

<sup>75</sup> As far as I can tell, deformed desires and adaptive preferences are used interchangeably in this debate.

The final account I consider is Anita Superson's strong substantive account that she presents in her paper "Deformed Desires and Informed Desire Tests" (Superson 2005). Superson finds desire-satisfaction theories of rational choice lacking since they fail to classify deformed desires as irrational. But such desires are irrational, she argues, since they undermine an agent's equal moral worth. The brief discussion about the classification of Stoljar and Charles' accounts is useful here as well, since Superson also mentions the impact that deformed desires have on an agent's normative competence. In Superson's case, it is an agent's ability to see herself as having equal worth that is compromised by deformed desires. But according to Superson, it is not just an agent's ability to see herself as worthy that is compromised; the agent's *objective* equal moral worth is compromised when she adopts deformed desires (Superson 2005: 122). Thus, the content of an agent's desires is again open to direct normative constraints.

### ***6.2. Deformed Desires, Revisited***

Recall that the purpose of this discussion is to extract a sufficiently detailed account of deformed desires to determine whether or not my criticism holds; that is, whether or not it is true that on strong substantive accounts of autonomy we must conclude that women are mere dupes of patriarchy. This would be an unwelcome result, since these very same accounts of autonomy are meant to vindicate Stoljar's Feminist Intuition. More broadly, these accounts are designed to yield normative results that are consistent with feminist aims. If it is the case that women are, on these accounts, bound to be dupes, I have argued that the result is in fact normatively unsound from the very same feminist perspective.

Perhaps it is the case, then, that there is a more detailed account of deformed desires that prevents this unwelcome result. Unfortunately, the exact relationship between deformed desires

and the denial of autonomy remains unclear. Consider again Stoljar's Feminist Intuition – that preferences *influenced by* oppressive norms of femininity cannot be autonomous. But she alternatively describes deformed desires as those “*motivated by* oppressive and misguided norms” (2000, 98) and “*derived from*” oppressive norms (Stoljar 2000: 109). Sonya Charles argues that desires *resulting from* false beliefs that rely on subordinating reasoning and perpetuate oppression should not count as autonomous (Charles 2010: 413). But she alternatively describes deformed desires as those that “*express*” oppressive norms of femininity (Charles 2010: 416), “*replicate*” oppressive norms (*ibid.*), and “*contribute to*” an oppressive system (Charles 2010: 417). Later, Charles lists the three conditions that are required for a desire to be deformed: they must represent false beliefs, they must rely on subordinating reasoning or lead to subordination, and they must perpetuate the agent's own oppression or perpetuate oppression in general (Charles 2010: 416). But surely relying on X and leading to X are not the same; replicating X is not the same as resulting from X, which is in turn different from expressing X. Are we meant to think that deformed desires result in all of these cases?

At the very least, this confusion highlights the need for clearer definition of deformed desires. But why might it also highlight a potential tension? Recall that part of the problem with deformed desires is that they benefit the oppressor rather than the oppressed. But constitutively relational accounts of autonomy which mandate that being in the wrong kinds of relationships – such as, for example, living in a patriarchal society – renders an agent definitionally non-autonomous may *themselves* benefit the oppressor rather than the oppressed.<sup>76</sup> Privileged agents,

---

<sup>76</sup> The benefit is expressed by the theorists themselves, who equate having autonomy with the possession of full agency (Superson 2005, 111, 117). The possession of full agency is also linked to being treated with respect (Nedelsky 2011, 65). There is obviously more to be said about the value of autonomy, but I will not discuss the topic here.



on these accounts, have autonomous desires while non-privileged agents do not. Men can be autonomous, but women can't.

This tension, I argue, reveals a feminist ambivalence about individualism. Note that, on the robustly relational accounts I have discussed, agents are required to free themselves from oppressive relationships in order to be autonomous. Now consider what we might call a *Social/Social* account, which expresses both **SST(2)** and **SAT(2)**, that is, a robustly relational conception of the self, and a robustly relational conception of autonomy.<sup>77</sup> Stoljar and Charles, for example, provide accounts that can be classified as *Social/Social*. While maintaining that autonomy is constituted by our social context – and while at the same time maintaining that we ourselves are constituted by our social context – these theories require us to separate ourselves *from this very same context* if we are to be autonomous. In other words, in order to be autonomous, we must both reject and embrace a type of individualism.

## **7. Relational, Reconsidered**

In this chapter I've presented arguments against conceiving of selves as socially constituted, and against conceiving of autonomy in the same way. What I've called *Social/Social* accounts of autonomy, which hold that both selves and autonomy are robustly relational, also require a certain level of individualism; that is, such accounts require agents to overcome, or think outside of, their oppressive circumstances in order to be autonomous. But this level of individualism is not possible unless selfhood and autonomy are less relational than is specified by such accounts. Perhaps what is needed is a push back toward individualism. This push doesn't force us back

---

<sup>77</sup> Thanks to Luke Semrau for this moniker.

into an atomistic, ahistorical, or hyper-individualist theory. Rather, it acknowledges that our social context influences both our self-conception and our autonomy, and that certain relationships might help or hinder our ability to be autonomous at all. We can accept these truths without reverting to social atomism, but we can also accept these truths without thinking that our social context wholly determines our identity or our autonomy. In other words, we can accept that we are not islands without admitting we are dupes.

## CHAPTER V

### PURELY PROCEDURAL FEMINIST AUTONOMY

In the previous chapter, I spelled out the problems I see with strong substantive versions of relational autonomy. In doing so, I attempted to switch the burden of proof from procedural theorists to proponents of strong substantive theories. In this chapter, I will discuss my own positive theory – a procedural account of relational autonomy.

#### **1. What We Want and What it Costs**

##### ***1.1. What Do We Want?***

It's important to ask what it is a theory of autonomy is supposed to accomplish. What do we want a theory of autonomy to do? Recently, relational autonomy theorists have started to ask this question more explicitly. Paul Benson, for example, argues that “future progress in work on relational autonomy [...] will depend in large measure on greater attention to what feminists really need a conception of personal autonomy to do in relation to the normative social purposes of confronting and overcoming women's subordination” (Benson 2014: 89). Joel Anderson, too, has recently suggested that “we could try to articulate competing conceptualizations of autonomy in terms of competing, more-or-less coherent packages regarding which ways of treating people are appropriate, what kind of society we want to live in, what should get priority, how we can arrange our society so as to increase our ability to realize certain values, etc.” (Anderson 2014: 356). I agree, and so I devote the first section of this chapter to answering this question.

I think the best way to proceed is to ask, first, what we *want* from a theory of autonomy, and then to balance these goals with the *costs* we are willing to pay. To help explain, let's look again at the feminist strong substantive accounts I considered in the previous chapter. I argued

that these theorists share one central goal – they want their theories to be able to account for the normative commitments of feminism; more specifically, they want theories that will address the problem of internalized oppression. Natalie Stoljar’s account of the Feminist Intuition is a quintessential example of this approach; her account explicitly prioritizes the problem of internalized oppression (Stoljar 2000).

But this prioritization comes at a cost, which I also addressed in the previous chapter. The cost, I argued, is that the choices of individual women who profess to hold an oppressive preference are not respected, and the women themselves are considered to have less than full agency (Superson 2005: 111).<sup>78</sup> In fact, Stoljar recognizes that her account is open to such a charge but ultimately finds it unconvincing (Stoljar 2014: 231). The important point here is that accounts like Stoljar’s have a clear goal – addressing internalized oppression – and are willing to pay the cost: in some cases, such accounts will be forced to conclude that the self-understandings of certain women are mistaken. I have attributed this cost to a focus on the social nature of personhood and autonomy.

I think this cost is too great, and so I attempt to tip the scale back in the other direction. How can we avoid the conclusion that a great many women lack full agency? I suggest that we can do so by prioritizing women’s voices – that is, by prioritizing the individual instead of the social. This is the first thing a theory of autonomy should accomplish; it should take an agent’s subjective perspective seriously. In order to do so, priority must be given to an agent’s self-understanding. In cases where this understanding reveals oppressive preferences, we should listen to the agent’s self report before jumping to conclusions based on the content of her

---

<sup>78</sup> Stoljar might disagree that her account denies full agency to women with oppressive preferences, but in order to do so she must distinguish between autonomy and agency. Although she argues that they are not equivalent concepts, she does not explain what distinguishes them (Stoljar 2014: 249).

preferences. The reasons for this prioritization are well rehearsed in the literature. Recent discussions have re-asserted familiar worries – that overriding women’s self-understandings of their own choices amounts to telling these women that they are confused about their own good (Meyers 2014b: 432), and that denying the rationality of their choices is at least sometimes descriptively inaccurate and always *prima facie* normatively troubling (Khader 2012: 305). There are also epistemological concerns. Women bargain with patriarchy in varied and complicated ways; sometimes this bargaining includes conforming to patriarchal norms.<sup>79</sup> Accounts that don’t prioritize women’s self-understandings risk cultural insensitivity, because they may fail to take account of these complexities (Sperry 2013: 892).<sup>80</sup>

It is for these reasons that I reject giving priority to addressing internalized oppression. Instead, I give priority to women’s own self-conceptions. Here is how such an account will compare to strong substantive accounts:

#### Strong Substantive – Social Focus

1. Address internalized oppression
2. Value women’s self-conceptions

#### Procedural – Individual Focus

1. Value women’s self-conceptions
2. Address internalized oppression

---

<sup>79</sup> Bargaining with patriarchy is also a phrase attributable to Narayan.

<sup>80</sup> I want to be clear that I’m not charging strong substantive accounts with *ignoring* women’s own self-conceptions. This objection would be misguided, since Stoljar, Superson, and Charles, for example, are all clearly concerned with women’s subjective perspectives. Ultimately, however, they conclude that deformed desires (or preferences formed under conditions of oppression) are not *really* the agent’s own in any significant sense, and so overriding them does not involve overriding *her* real preferences at all (e.g. Superson 2005: 115; Stoljar 2014: 228).

I understand these orderings as lexical priority rankings,<sup>81</sup> according to which we must consider the first condition before moving on to consider the second. This is, at this point, severely under-explained. As the chapter progresses, I will introduce an example to illuminate what I have in mind.

### ***1.2. What Does it Cost?***

As I said above, there are two questions we should ask ourselves when constructing a theory of autonomy: what do we want a theory of autonomy to accomplish, and what are the costs we are willing to pay? I argued in the previous chapter that the cost of feminist strong substantive theories is that many women will be construed as lacking full agency. But there are costs to my general approach as well. Because my theory is broadly procedural – that is, it does not place direct normative constraints on the desires of agents – it is often argued that it cannot adequately take account of internalized oppression (e.g. Stoljar 2000: 95; Charles 2010: 413). And indeed, since I have now explicitly stated that I give women’s self-conceptions priority over addressing internalized oppression, it might be argued that my theory egregiously fails to address the oppression of women.

I will fully address this objection after filling in the details of my account. For now, I want to acknowledge that this objection is the biggest hurdle for my theory, and there will be some fairly large bullets to bite. However, I also think that there is a way to understand what I’m doing as putting the worry about oppression at the forefront. Here, I direct attention back to Chapter One, and my discussion of political liberalism. I discussed Rawls’ motivating thought:

---

<sup>81</sup> See John Rawls’ *Theory of Justice*. A lexical priority ranking is “an order which requires us to satisfy the first principle in the ordering before we can move on to the second, the second before we consider the third, and so on. A principle does not come into play until those previous to it are either fully met or do not apply” (37-38).

that any attempt to justify state coercion by appeal to a specific moral doctrine would ultimately fail to respect all citizens as free and equal. He thought that any attempt to establish liberalism by appealing to a specific moral doctrine would require forcing this doctrine on other citizens who adopt a moral doctrine different from comprehensive liberalism. For Rawls, this would constitute a type of oppression (Rawls 1993: 9-10).

A version of this same thought motivates my content-neutrality. Imagine the Deferential Wife, for example. Most of her commitments look problematic from a feminist perspective; that is to say, most of her commitments look problematic from the particular moral space we occupy as feminist theorists. But to deny that this woman's life choices are autonomous is to deny that she has full agency – and this is the case *even* according to the strong substantive theorists I've discussed.<sup>82</sup> Superson argues, for example, that “deformed desires jeopardize autonomy *and thus full agency*, since their satisfaction aims to lower their bearer's value” (Superson 2005: 111). In later work, she maintains that desires which are “deformed by social circumstances and stand in the way of [a person] being fully autonomous” compromise this person's agency (Superson 2010: 254). But to deny full agency is often to welcome paternalistic intervention by others who hold the ‘right’ moral view, or by the state. And this intervention, in turn, may itself be the kind of oppression that John Rawls imagines in *Political Liberalism*, that is, the forced adoption of a particular moral view that differs from one's own. Forcing the Deferential Wife to

---

<sup>82</sup> Procedural theorists often stress this point. See, for example, Diana Meyers: “Restrictive, value-saturated accounts of autonomy are troubling because they promiscuously stigmatize women as victims and because they homogenize authentic selves and autonomous lives” (Meyers 2002: 16). Charles acknowledges that this charge is leveled at strong substantive theories, but concludes that “the implications of this are less far-reaching than procedural theorists believe” (Charles 2010: 427). But I don't think *mentioning* that the implications are less far-reaching does enough to *demonstrate* that fact.

adopt a different moral perspective *because we (think we) know better* will look a lot like oppression to her. And this at the very least should give us pause.

### ***1.3. Feminist Husbands and Dominant Wives***

So far I've just illustrated the negative practical consequences of feminist strong substantive accounts, but there are also negative theoretical consequences. Insofar as we want a theory of autonomy to be extensionally accurate, feminist strong substantive accounts will run into trouble. To explain what I have in mind, I offer two examples.

**Feminist Husband:** Imagine a feminist philosopher named Maribeth who, by all external appearances, lives her life according to feminist principles. At home, she and her husband Michael split all tasks evenly. She does the dishes, but he does the laundry; she packs lunches for her children, but he drives them to soccer practice, etc. But it is Michael, and not Maribeth herself, who is committed to women's equality. He is a strong believer in female equality, and knows that, were Maribeth left to her own devices, she would prefer to be subservient to men. Michael capitalizes on Maribeth's desires for subservience, and commands that she live her life according to the feminist principles to which he is so devoted. Maribeth obeys, because she believes that wives should always do as their husbands command. Both are completely satisfied with their marriage.

**Dominant Wife:** Imagine a feminist lawyer named Linda who is committed, in all aspects of her life, to ending women's oppression. She believes that, until our society radically changes, the only way to achieve this goal is for women to be dominant over men. Her husband, Lloyd, has



never really been committed to anything, and so goes along with whatever Linda says. She pays all the bills, works outside the home, and does all the traditionally “male” tasks, while Lloyd stays at home, cleans the house, etc. Lloyd cooks Linda’s favorite meals for dinner, dresses in Linda-approved clothes, and defers to her judgment when making decisions. The house is in her name, the bank account is hers alone, and she leaves Lloyd a small allowance each week to spend on groceries and frivolities. Linda, however, was raised by a domineering mother who forced her to have feminist views. She has never reflected on her desires, but takes them as given. She is only committed to ending women’s oppression because of a fear of disappointing her mother. Linda and Lloyd are both satisfied with their marriage.

Both Maribeth and Linda have desires with the right sort of content, according to the feminist strong substantive accounts I considered in the previous chapter. Both have positive desires to end women’s oppression, and both appear to be committed to their own equality (or in Linda’s case, her own dominance). On these accounts, Maribeth and Linda would be considered autonomous. Their desires were not formed under conditions of oppression, and they do not align with oppressive and stereotypical norms of femininity. But there are complexities in both cases which should make us question their autonomy.

Consider Maribeth. Her desires are entirely other-directed. The only reason she appears to be committed to feminist goals is because she is *really* committed to her own subservience. Because her desires align perfectly with feminist desires, they do not raise the Feminist Intuition. She is still unflinchingly subservient to a man, however, so perhaps I was too quick in concluding that feminist strong substantive accounts would find her autonomous. But the only way for them to reach the opposite conclusion would be to prioritize her second-order desire for

subservience over her first-order feminist desires. On what principled, non-arbitrary grounds could this be done? Linda's case is even more complicated. She is not subservient to a man; she is so dominant in all aspects of her life that it would be difficult to call her subservient at all. One might be able to argue that she is subservient to her mother, but this does not raise our Feminist Intuition.

However, I argue that we do have reason to question Maribeth and Linda's autonomy. Both of their desire sets, although aligned with feminist commitments, are entirely other-directed. Maribeth longs to please her husband, and Linda is motivated by a fear of disappointing her mother. There is a sense in which their desires do not belong to them. It's important to note that I am not begging the question here; all I am trying to do is assess our intuitions. Feminist strong substantive accounts give us a theoretical chainsaw when we need a scalpel. Elizabeth Sperry points this out when she argues that, on feminist strong substantive accounts, "suspicious decisions made by nonoppressed agents are nevertheless called autonomous, whereas suspicious decisions made by oppressed agents are called heteronomous" (Sperry 2013: 899). It seems like there is much more to autonomy than just the simple alignment of desires with feminist commitments, and insofar as these accounts cannot recognize such contextual complications, they may fail to be extensionally accurate.

## **2. Agent Centered, Liberty Oriented**

So far I've discussed the motivations behind my account and offered a basic outline; but I've been painting with very broad brushstrokes. In this section, I will spell out a more detailed version of my account. Note that I am focused here on authenticity conditions. Agents will also meet competence conditions, according to which they are able to engage in the process of critical

reflection. But here I concentrate on what it takes for a desire to belong to an agent, so that we can properly say that it is hers. How do we know when a desire or preference is authentic? We can then determine if the actions which arise from this desire are autonomous.

### ***2.1. Reflective Endorsement***

Reflective endorsement is a key element of procedural accounts. Basic accounts of reflective endorsement include the idea that an agent must somehow bring her desires into view and then reflect upon them. This reflection will include asking herself whether or not she endorses the particular desire in question, or whether or not she identifies with it, or – according to Frankfurt, for example – whether or not she wholeheartedly identifies with it (Frankfurt 1988). If, upon reflection, she endorses or identifies with the desire, then action according to this desire is autonomous action.<sup>83</sup> If, on the other hand, she disapproves of the desire or finds that she feels alienated from it, then action according to this desire is heteronomous action. This is perhaps the simplest account of reflective endorsement, and it has been subject to well-known criticisms (which I discussed in Chapter Two).

The reason some version of this account of reflective endorsement is utilized by procedural theorists is because such endorsement is meant to signal *authenticity* (e.g. Meyers 1987: 619). Autonomous action is self-guided; in order for an agent's actions to be autonomous, the 'self' guiding her actions must be *hers*. The problem with very basic accounts of reflective endorsement is that they leave the door open for manipulation, since the endorsement itself

---

<sup>83</sup> By 'according to' here, I mean that the desire must contribute causally to the action in question (e.g. Friedman 2003: 5). Michael Bratman refers to this phenomenon as an "effective motive of intentional action" which he defines: "When a desire for E so functions, it is – at the least, and roughly speaking – part of a mechanism that cognitively tracks and thereby tends to promote (given true beliefs) bringing about or realizing E" (Bratman 2002: 66).

might be externally influenced. In other words, it is possible for someone to reflect on a particular desire, endorse it, and then act on it – without this desire really belonging to him in any real sense; it could be the product of internalized oppression, for example.

I propose, however, that some account of reflective endorsement must be central to any procedural theory of autonomy. The task at hand, then, is to develop an account of reflective endorsement that is not so obviously open to criticism. I will do this in two ways: first, I will develop a more detailed account of reflective endorsement as such; and second, I will supplement reflective endorsement with other conditions that are necessary for autonomy. The account, briefly, is as follows:

1. Reflective Endorsement (RE): An agent (A) reflectively endorses a desire (D) when she considers – to some extent – whether or not D coheres with the greater set of values (V) that constitute her identity.
2. Diachronic Value Cohesion and Narratability: Either D coheres with V over time, and/or A is willing and able to narrate how she came to hold D and how it relates to V.

There are two things to note. First, RE requires actual, and not hypothetical, endorsement – that is, A must *actually* examine D and whether or not it coheres with V. The immediate worry is that we do not, in practice, examine many of our desires before we act; this worry inspires some procedural theorists to adopt a hypothetical account of endorsement, where an agent reflectively endorses a desire if she *would* endorse it, *were* she to reflect upon it. Hypothetical endorsement, though, opens the door for a problematic sort of externalism; it might allow consideration of an agent's reflection from a standpoint other than her own, indeed, even from a

standpoint that she rejects.<sup>84</sup> Instead of proposing a hypothetical account of reflection, then, I endorse an actual account that admits of degrees (Meyers 1987; Friedman 2003). Reflection can be more or less robust. In other words, if A acts on D and feels immediate regret, she might ask herself why this feeling occurred. This questioning is a minimally robust sort of reflection, which might lead A to conclude that she does not, in fact, endorse D. Note, too, that this reflection was sparked by something that A *felt*. There is no reason that emotions or feelings (like the feeling of regret) couldn't initiate a process of critical reflection. I also draw the reader's attention back to the Introduction, where I discussed the *types* of desires and actions I would be considering. Although cases about Brownie Batter ice cream are illustrative, these are not the kind of case I am principally concerned with. I am instead interested in more impactful, or major, decisions that will help direct the course of an agent's life.

Second, the coherence at the core of reflective endorsement may often be signified by feelings of nonalienation, rather than positive feelings in favor of D.<sup>85</sup> I understand alienation, following Christman, as more than just a lack of positive feeling, but rather as involving a feeling of constraint, or of wanting to rid oneself of the desire in question (Christman 2009: 143). Accounts like Frankfurt's which require a positive assessment of D (like wholehearted identification) require too much, since we may often feel ambivalent. John Christman notes that, were autonomy to require positive valuations of our desires, then "only the supremely lucky and fulfilled among us would count as autonomous" (Christman 2009: 13). And, for reasons that I discussed above, we have good reasons to reject a conception of autonomy that reserves autonomy as an ideal to be achieved only by the lucky few.

---

<sup>84</sup> See my discussion of Christman in Chapter Two.

<sup>85</sup> See Christman 2009. Although, of course, she *may* have positive feelings of identification, such feelings are not required.

To fill in my conception of reflective endorsement with sufficient detail:

An agent (A), *perhaps sparked by feeling or emotion*, reflectively endorses a desire (D) when she *actually* considers – to some extent (*perhaps very minimally*) – whether or not D coheres with the greater set of values (V) that constitutes her identity, *where this coherence is signified (minimally) by a lack of alienation*.

Described in this way, this account of reflective endorsement provides a foundation for the authenticity that is necessary for autonomy.

## ***2.2. Diachronic Value Cohesion***

But what exactly is this ‘greater set of desires’ that constitutes our identity, and what does it mean for a specific desire to ‘cohere’ with it? First, let me start by explaining my motivation for introducing this particular aspect. Critics of procedural theorists have pointed out that reflective endorsement alone does not adequately screen for the pernicious influence of internalized oppression.<sup>86</sup> I agree that, while some sort of reflective endorsement is necessary for autonomy, it is certainly not sufficient. Why? Because it looks at least plausible that an agent in question has reflectively endorsed a desire or preference that aligns with oppressive norms.

In the example, the Deferential Wife (DW) demonstrates that she (again, at least plausibly) thought about her desires and the actions that they motivated, and she cites reasons in its favor – reasons that, from her perspective, are *decisively* convincing. This example illustrates what critics of procedural accounts point out – that reflective endorsement can itself be the

---

<sup>86</sup> See Chapter Four.

product of oppressive circumstances. It is easy to tell a story about how a patriarchal society or a patriarchal religion, in essence, formed DW's preferences for her. And while I think it is possible, in some cases, that the process of reflection might weed out oppressive preferences, the story of DW demonstrates that this won't always be the case.

There are at least two ways to combat this weakness. First, DW could consider whether the values she acted upon were consistent with her other recognized values, and second, she could consider her value set across a longer period of time instead of just at a moment. This latter point is motivated by the fact that we exist diachronically, that is, across time (Christman 2009). To determine if a desire is self-governed, then, we shouldn't simply consider it in relation to the values that we hold at a particular time. Instead, we should consider the desire in relation to the values that we have consistently held over a longer period of time. This helps to ensure that the authenticity we are concerned with is representative of who we are, full stop, and not just who we are *today*.

This is a colloquial way of putting the point, but here's a very simple case that illustrates what I have in mind: remember Betsy, from Chapter Two, who had a strong desire to eat an entire pint of ice cream. If Betsy considers this desire only in relation to the values she holds at this moment, she would easily conclude that the action is autonomous. Why? Because the values she holds at this moment are just those values that are motivating her to act: she values immediate satisfaction, maybe, or the feeling of treating herself to something special.<sup>87</sup> But if she were to consider the desire in relation to the values she has consistently held over a longer period of time, she might reach a very different conclusion. Maybe she values health over immediate satisfaction, and the feeling of respecting her body over the feeling of receiving a

---

<sup>87</sup> Note that this won't always be the case – Betsy might be suffering from weakness of will.

special treat. Something like this is arguably what happens when Betsy pauses at the freezer and concludes that she doesn't *really* want the ice cream after all.

To return to DW, this *diachronic value cohesion* might help demonstrate her lack of autonomy. The values she considers at this moment might include fulfilling the needs of her husband. But if DW were to feel a deep dissatisfaction with her lifestyle, this feeling might inspire her to reflect on her desire, and this reflection might help her realize that being a submissive wife is not, in fact, consistent with the values she has consistently held across her lifetime. Perhaps she values independence, feeling like a capable woman, and putting her own needs – just sometimes – above the needs of others. Remaining deferential might not reflect who she is, and thus making this choice would be non-autonomous.<sup>88</sup>

It's important to note, though, that diachronic value cohesion is not just a convenient or *ad hoc* requirement designed to define all cases of chosen deference as non-autonomous. Although it is possible that this requirement could rule out deference in this way, it is also possible that such preferences could survive the test. DW could also have long-term values that favor deference, so that acting upon this desire would cohere with her long-term value commitments. In this case, we may be able to conclude that she is autonomous. I will say more about this in the following section.

---

<sup>88</sup> Of course, it is also possible that she won't feel any dissatisfaction, and she won't critically reflect on her desires and values *at all* (even though this is a minimal requirement). In this case, we would also question her autonomy. However, if she were to reflect upon her desires and then endorse them, and these desires were consistent with her values, I think we have reason to question her non-autonomy. I will discuss this more in Section 3.



### 2.3. *Narratability*

So far, then, I have proposed two conditions for autonomy: reflective endorsement and diachronic value cohesion. But two objections loom: first, can't people change? And second, can't people's consistent long-term values be non-autonomous? The first objection questions whether or not value consistency over time is *necessary* for autonomy, and the second questions whether or not value consistency is *sufficient*. Regarding the first charge, maybe it is the case that DW reflects upon her desires and finds that she has always valued being deferential. If she then considers whether to leave her marriage because she feels that she desires independence, her desire might look like it appears from nowhere; it does not cohere with the set of values she has consistently held over a long period of time. It is a new value. Thus, although value consistency looks like a promising condition for autonomy, we have at least one reason for thinking it is not necessary for autonomy.

I therefore propose the concept of *narratability*: if an agent suddenly finds herself endorsing drastically different desires, she must be able to tell a coherent story about her shift in values. I am inspired here by Andrea Westlund's account of answerability, although there are aspects of her theory with which I disagree (Westlund 2003, 2009, 2012). What follows is a brief discussion of Westlund's view, after which I will describe my own account and the important ways it differs from Westlund's. I'll then move on to consider the second objection, regarding the sufficiency of value consistency for autonomy.

Westlund is concerned with cases similar to the case of DW – cases where an agent is deferent or submissive to such a degree that we begin to wonder whether her desires are really *hers* at all. Westlund argues that the reason women like DW appear non-autonomous to us is because they fail an answerability criterion; they fail to hold themselves answerable to external

critics who question the authenticity of their desires (Westlund 2003: 485). She argues, “Autonomy requires a disposition to hold oneself answerable to external, critical perspectives on one’s motives and desires, because such a disposition marks the relevant distinction between being gripped by and governing the practical reasoning that guides one’s actions” (Westlund 2012: 65). The intuitive idea is that, if we are unable to answer critics regarding why we have a particular desire, it is likely that the desire is *governing us* rather than the other way around. Failing to meet the answerability criterion is thus failing the test of authenticity.

There are a few crucial aspects of Westlund’s theory to note: first, she is only concerned with the *disposition* to hold oneself answerable to external critics, and not with actual engagement with them. All that matters is that an agent sees herself as willing and able to respond to critics who question the authenticity of her desires (Westlund 2009: 35). Second, Westlund understands this disposition as fundamentally relational. She argues that “having such a disposition means positioning oneself as always a potential member of a reflective or deliberative dyad, which is one aspect of the relationality of the view” (ibid.). The view is relational also because the autonomous agent holds herself responsible for her desires; she must be able to “represent herself in interpersonal dialogue” (ibid.), even if this dialogue occurs within her own imagination or with non-actual others (Westlund 2009: 36). Westlund thus proposes a constitutively relational account of autonomy that is nonetheless procedural.

I agree that there is something about being willing and able to explain a desire that signals autonomy. But I’m not sure why we have to be willing and able to answer to *external* critics. Don’t we often engage in dialogue with ourselves? In fact, I think we can understand reflective endorsement in such a way that this process is itself a sort of dialogue between ourselves and our desires. Westlund admits that we can dialogically engage with *imaginary*

others, but insists that this this is still a type of engagement with external critics. “Either way,” she argues, “responsibility for self depends on the internalization of a very basic sort of interpersonal relation – namely, a form of justificatory dialogue that (presumably) we begin to learn in our early interactions with parents and other caregivers and continue to develop throughout the process of maturation” (Westlund 2009: 36).

But this, at best, demonstrates that dialogue is a skill we *originally learn* in social environments – it might be causally relational, but not constitutively so. In other words, it might be the case that we first learn to answer criticism that comes from others before we internalize critical dialogue. After we have internalized it, however, we don’t need to engage with others to subject our desires to critical scrutiny. This would mean, in turn, that the process of engaging in critical dialogue is causally relational, that is, our ability to so engage arises from our relations with others in much the same way that our identities, themselves, are formed. But the account would no longer be constitutively relational, since our relations with others don’t continue to constitute the dialogical process itself.

Instead of proposing answerability like Westlund, which includes a relational connotation – you are answerable *to others* – I propose narratability. That is, if we are suddenly faced with a new desire we should be able to *narrate* the values that led us down this road; this narration could include a dialogue with external others, but it could also include explaining to *oneself* how this desire arose. Narratability is thus not constitutively relational like answerability, that is, it does not require ongoing relationality.

I do agree with Westlund, though, that this aspect of autonomy should be hypothetical: an agent must have the *ability* to narrate how she came to value something new (where this ability includes a willingness to engage in this narration), but she needn’t *actually* do so. Why?

Because actual narration would be an extremely demanding requirement. Many of our deepest values are given to us by our social circumstance. We can of course come to endorse them, but this is far easier than narrating why we hold them so dear. Requiring that we tell a consistent story about each of our desires and values would be much too strict a requirement. All that is required is that we could tell such a story. And the need for this requirement will only arise in cases where an agent is faced with a new desire which seems to appear from nowhere but which *feels right* to her. The narratability criterion thus gives my theory the ability to account for personal growth and change over time, something which my account might otherwise lack.

I have yet to address the second objection noted above, regarding whether or not diachronic value consistency is really sufficient for autonomy. What do we say about cases in which an agent's consistent long-term values appear non-autonomous? Although narratability is helpful in cases where an agent finds herself endorsing a new value, suddenly out of nowhere, I also think it is helpful in cases where *no change* of value is noticed. I propose that narratability is thus required for autonomy, even in cases where an agent has a consistent set of values over time. The idea is that oppressive preferences might be weeded out by the combination of reflective endorsement, diachronic value cohesion, and narratability. Of course, it is still possible that an agent could reflectively endorse a desire that coheres with her broader set of values over time, and she is able to narrate how this particular desire relates to her broader value set. In these cases, my account will conclude that she is autonomous. In the following section, I will examine a hard case in order to further explain my account.

### 3. A Hard Case: There is No ‘You’ in Quivering

Combining the conditions required for autonomy explained in the previous section yields the following account:

An agent A is autonomous relative to a desire D when the following authenticity conditions are met:

1. Reflective Endorsement: A reflectively endorses D, where this endorsement consists in actually considering – to some extent (perhaps very minimally) – whether or not D coheres with the greater set of values (V) that constitutes her identity, where this coherence is signified (minimally) by a lack of alienation.
2. Diachronic Value Cohesion and Narratability:
  - a. *Either* D coheres with V over time AND when D does cohere with V, A is willing and able to narrate (to herself or external others) how D relates to V.
  - b. *Or*, in cases where D does not obviously cohere with V, A is willing and able to narrate (to herself or external others) how she came to hold D and how it relates to V.

A brief point of clarification is necessary at this point. In the previous section, I provided two justifications for including narratability in my account. The first justification is that agents must be able to remain autonomous through identity-changing experiences. In cases where an agent fails to meet diachronic value cohesion, that is, in cases where a new value appears as if from nowhere, the agent’s desires will no longer exhibit cohesion. *Instead* of meeting this condition, I propose that she could move on to the narratability condition. If she is willing and able to

narrate to herself (or others) how she came to hold the desire in question and how it relates to her broader set of values, then she is autonomous even though she fails the first part of the second condition.

The second justification for including narratability is that reflective endorsement and diachronic value cohesion do not collectively seem enough to rule out obvious failures of autonomy. Thus, even in cases where reflective endorsement and diachronic value cohesion are met, an agent *must still satisfy* the narratability condition. According to my account, then, reflective endorsement and narratability are necessary for autonomy. Diachronic value cohesion, however, is not always necessary. Failing to meet this condition does not automatically *disqualify* a desire as non-autonomous as long as the narratability condition is still met. But meeting the diachronic value cohesion condition does not automatically *qualify* a desire as autonomous; the narratability condition is required in these cases as well.

At this point, it is helpful to consider an example.<sup>89</sup> For the past thirty years, a movement of fundamentalist evangelical Christians has been burgeoning in the United States. This movement is inspired by two major sources: Mary Pride's 1985 book, *The Way Home: Beyond Feminism, Back to Reality*, and Psalm 127, read literally – “Lo, children are an heritage of the Lord: and the fruit of the womb is his reward. As arrows are in the hand of a mighty man; so are children of the youth. *Happy is the man that hath his quiver full of them*” (Psalm 127:3-5). As a result of this literal reading, Quiverfull families condemn the use of birth control (including traditional family planning measures), and praise the submission of wives to their husbands – as mandated by the verse: “For the husband is the head of the wife even as Christ is the head of the

---

<sup>89</sup> Following feminist writers like Oshana (2006) and Stoljar (2000) who consider feminist “hard cases” to illuminate their accounts. I have chosen the example purposefully to combat a tendency to disproportionately ascribe autonomy-failings to women in the Global South.

husbands” (Ephesians 5:23-4). Quiverfulls have a variety of other commitments similarly inspired by a literal reading of the Bible: women must stay at home and homeschool their children, and they must dress modestly and forego all economic independence.<sup>90</sup>

Even more sinister, though, are those commitments which arise from a literal reading of verses like the following: “And Adam was not the one deceived; it was the woman who was deceived and became a sinner” (1 Timothy 2:14-15); “Whosoever therefore resisteth the power, resisteth the ordinance of God: and they that resist shall receive to themselves damnation” (Romans 13:1-2); and finally “For his anger is but a moment, and his favor is for a lifetime” (Psalm 30:5). The result is that women are isolated in the home without access to economic resources; they submit themselves to their husband’s will, even when this will is violent, because their husbands (just like God) might occasionally express their love through violence; and they see child-bearing as the way to cleanse their souls of the guilt over causing the downfall of Man (my choice of words here is deliberate).

And yet, it is the Quiverfull women who are the movement’s greatest advocates. For example, a popular Quiverfull website, *Above Rubies*, is run by a woman – Nancy Campbell – who offers encouragement for other Quiverfull women. “Are you feeling base and discouraged? Don’t listen to those lies any longer. Lift up your head and embrace your mandate from God. You are not working for any earthly employer, but for the King of kings and Lord of lords, the Sovereign God of the universe.”<sup>91</sup> One of Campbell’s most popular blog posts, “Top 10 Time Wasters for Women” instructs, “Thinking about yourself is a deadly time-waster. You can’t

---

<sup>90</sup> I see the example of Quiverfull women as a real-life case of a Deferential Wife (Hill 1991) or Angel in the House (Westlund 2003; Oshana 2006).

<sup>91</sup> Thursday, February 12, 2015.

adequately nurture your children or bless others while you use all your emotional energy on yourself.” The post has over 24,000 hits.

So what do we say about Nancy Campbell and other Quiverfull women? Are they mere dupes of patriarchy? It seems clear that the values espoused by the Quiverfulls are just the repackaged values of a patriarchal society. Anita Superson would almost certainly conclude that these women, like Muslim women who veil, are not autonomous agents because they “lack sufficient belief in [their] self-worth” (Superson 2005: 115). The desires of Quiverfull women (to stay at home, to submit to their husbands, to forego economic independence) are likely deformed, because they benefit a patriarchal society and a patriarchal religion, but (arguably) do not benefit the agent herself. Superson argues: “The woman who wants to be self-sacrificing because she believes that women have a duty to serve their families has a deformed desire that benefits the system, not herself” (Superson 2005: 116). Because the Quiverfull’s desires are deformed, they are not her own, and are thus non-autonomous. Natalie Stoljar, who has recently argued that, “Many of the adaptive preferences of concern to feminists are autonomy impairments” would agree with this conclusion (Stoljar 2014: 252).

Proponents of strong substantive accounts reach this conclusion because they construct accounts which give lexical priority to addressing internalized oppression. Note that lexical priority is not the same as temporal priority, so that on their accounts we might first ask ourselves, “What does the Quiverfull woman desire?” before moving on to ask, “Does this desire indicate the presence of internalized oppression?” From a strong substantive feminist viewpoint, an affirmative answer to the second question can *override* the answer to the first. But it might seem like these women profess a *very real desire* to live a Quiverfull lifestyle. This should dissuade us from concluding (at least so quickly) that they are non-autonomous.



The case of Vyckie Garrison raises exactly these issues. Vyckie Garrison is a formerly prominent member of the Quiverfull community who has since renounced this lifestyle. But consider her life as a Quiverfull woman. She and her husband jointly publish *The Nebraska Family Times* – a Quiverfull newspaper – and she lives according to the dictates of the movement for sixteen years.<sup>92</sup> She explains that she “proudly embrace[s] the ideal of biblical headship and submission” and “believe[s], as the Bible teaches, that it is the man who is ultimately responsible for the spiritual well-being of his wife and children.”<sup>93</sup> She describes her husband as a “loving, godly patriarch,” and justifies his position of authority: “I wouldn’t say that my husband *use[s] male privilege* to control and dominate me and the kids. Male privilege [is] his rightful position.”<sup>94</sup> She welcomes child-bearing as her chance for salvation. She cherishes her marriage to her husband, and relishes her marriage to God.

I don’t think it would be at all controversial to say that the *substance* of Garrison’s desires is problematic from a feminist perspective. It is on these grounds that feminist strong substantive theorists would deny her autonomy.<sup>95</sup> I agree with them about the problematic nature of Garrison’s desires, but disagree that this (alone) renders her non-autonomous. It is possible that Garrison reflects on her desires and endorses them, especially keeping in mind that endorsement is a very minimal condition. At one point she describes that it is she, and not her husband, who is the real supporter of patriarchy: “I *willingly* [go] along with all the harsh

---

<sup>92</sup> See Vyckie Garrison’s recent article: “How Playing Good Christian Housewife Almost Killed Me,” on a website devoted to helping women escaping the Quiverfull movement: *No Longer Quivering*. (Their tagline is worth mentioning, because it highlights the relevance of this case to my discussion of autonomy: “There is no ‘you’ in quivering...*No Longer Quivering* is a gathering place for women escaping and healing from spiritual abuse.”) Published September 19, 2014. <http://www.patheos.com/blogs/nolongerquivering/2014/9/vyckie-garrison-how-playing-good-christian-wife-almost-killed-me/>

<sup>93</sup> All quotes from Garrison are from this article.

<sup>94</sup> Her emphasis.

<sup>95</sup> Of course, the story isn’t that simple. See Chapter Four.

demands of the Quiverfull lifestyle, and in many instances, *I [am] the one* who pushe[s] patriarchy and headship ON HIM.”<sup>96</sup> Garrison professes a very real desire to live a Quiverfull lifestyle, and seems to provide a real-world example of Westlund’s hypothetical autonomous fundamentalist: “But if (*ex hypothesi*) a fundamentalist woman does freely and authentically accept a condition of social and personal subordination, it seems [...] problematic to assume that her condition as subordinate, in and of itself, undermines her status as self-governing agent” (Westlund 2009: 29).<sup>97</sup>

Now, as with most similar cases, the discussion becomes difficult, since we are trying to determine whether or not something internal to Garrison amounts to autonomy. This doesn’t mean that we should abandon all hope, it just means that we must examine several different outcomes to compensate for our uncertainty. One possibility is that Garrison has never reflected upon her desire for submission in even the minimal sense required. She never even gets to the point of considering this desire in relation to her deeply held values, but were she to ask herself about this desire, she would not be able to tell a consistent story to justify its role in her life. We can imagine her saying, with exasperation: “Why do I submit to my husband? Because...well because he’s my husband!” There is a sense in which this question doesn’t even make sense to her (Westlund 2003: 488). Perhaps she follows Nancy Campbell’s advice to Quiverfull women and doesn’t consider *herself* at all. She has not critically reflected on this desire; she is blindly devoted to her husband and to the movement. It is true that, from Garrison’s own account, she doesn’t seem like a blind devotee – but it is possible; in this case, most accounts (including my own) would conclude that she is not autonomous relative to her desire for submission. Action based on this desire would thus be heteronomous.

---

<sup>96</sup> Her emphasis.

<sup>97</sup> Westlund’s example is in response to Oshana’s *Taliban Woman* (2006).

But the reason this is a *hard* case is because of scenarios like the following: Garrison is a self-reflective person. She often considers her desire for submission, and each time she reflects upon it she endorses it. Perhaps she thinks that morality demands it of her, or perhaps she has internalized her husband's desires to such an extent that she no longer separates his desires from her own. Either way, let's stipulate that the reflective endorsement criterion has been met.

As is often the case, much depends on the specifics of the case. As it turns out, Garrison comes to feel a deep dissatisfaction with her life. She describes suffering "a complete mental and physical breakdown" as she realizes that her home "ha[s] become an oppressive, miserable place in which none of us [are] happy, and it fe[els] like we [are] all losing our minds".<sup>98</sup> This revelation causes her to rethink her desire for submission, and realize that she has a set of deeply held values that are inconsistent with it. What this might tell us is that, though Garrison has previously reflectively endorsed her submission, she has never thought to consider it in light of her other deeply held values. Were she to try to tell a consistent story about how this desire fit with these values, she would not be able to do so. On my account, and on strong substantive accounts, Garrison's desire for submission is non-autonomous.

But it's also possible that the desire was fully autonomous. We can imagine a different woman, perhaps Nancy Campbell, who has a different diachronic set of values than Garrison. Her deeply held values perfectly cohere with her desire for submission to her husband and to God. She can tell a consistent story about growing up in the church, and how God's love transformed her into the type of woman she always knew she wanted to be. (In fact, she makes similar claims on her website.) In this case, *even though we are talking about the same desire*

---

<sup>98</sup> Again, all quotes come from Garrison's 2014 article.

*for submission*, the details of the case demonstrate that Campbell (unlike Garrison) is autonomous regarding her desire for submission.

This hard case illuminates a few things about my account: first, we cannot conclude, just based on the content of peoples' desires, that they are or are not autonomous; much will depend on the value set of the person we are talking about. Second, reflective endorsement alone is sometimes not enough to weed out desires that are problematic for the agent herself. But perhaps most importantly, this discussion alludes to some fairly serious looming objections. I consider these in the next section.

#### **4. Biting Bullets?**

In this section, I will address looming objections, all of which revolve around the seemingly individualistic nature of my account. The focus of this chapter has been on reflective endorsement and the authenticity conditions necessary for autonomous agency. Some feminist theorists, though, argue that authenticity should not play such a central role; I will focus on Marina Oshana's argument to this effect. Oshana defines authenticity in the following way: "A person's cognitive, affective, valuational, and dispositional states, as well as personal commitments, social roles, and ideals are authentic if the person would 'wholeheartedly identify' with them or would embrace them without reservation were she to critically reflect upon their content and origin" (Oshana 2007: 413). She denies, though, that autonomous agents have to be authentic; it is perfectly compatible with autonomy for a person to be deeply dissatisfied with her values and desires. For Oshana, what matters for autonomy is that the person in question has *de facto* control over aspects of her life. "Among the indispensable components of autonomy," she

argues, “are substantive external social conditions. To be autonomous is to stand in a certain position of authority over one’s life, notably with respect to others” (Oshana 2007: 421).

Simply put, for Oshana, autonomy is not about authenticity.<sup>99</sup> It does not matter for Oshana how an agent relates to her own desires and values; identifying with her own self-conception is not important.<sup>100</sup> All that matters for autonomy is that an agent is a part of the *right kind* of social relationships, that is, relationships that allow her to exercise control over her life (Oshana 2007: 421) – and it can be objectively determined which sorts of relationships are conducive to or destructive of autonomy (Oshana 2005: 92). This is the core of her social-relational account of autonomy. Internalist accounts fail, for Oshana, because they are purely psychological – they determine an agent’s autonomy status by referring only to her psychological condition (Oshana 2006: 50). But because autonomy depends on our *de facto* control over social circumstances, our “sensation and feeling” of autonomy might be mistaken (Oshana 2006: 50). Oshana would take issue with my account because it places heavy emphasis on exactly the criterion she thinks is often irrelevant: it is internalist and subjective.

In other words, according to Oshana, my account does not place enough emphasis on the way that social forces can impact a person’s autonomy. This is closely related to a broader

---

<sup>99</sup> At times, Oshana says that authenticity *alone* cannot determine autonomy, making it sound like we need authenticity in addition to other criteria: “In addition to whatever psychological or dispositional characteristics are required for autonomy, and in addition to whatever can be said of a person’s psychological history, there are social criteria according to which we judge someone as autonomous” (2006: 50-1). But her discussion of case studies makes it clear, I think, that a person’s subjective experience of her own autonomy is often *irrelevant* for judging her to be autonomous (2006: 52-68). For example, of the Taliban Woman who submits herself to her husband and male relatives, Oshana argues: “Despite her authentic, competent, and sober acceptance of such situations...her lack of relational autonomy means that she is not self-governing” (2006: 100).

<sup>100</sup> This is why, for example, the female CEO who feels conflicted and has a frustrated will – she wants to stop having control over her life, and she would rather have someone else take control for her – is still autonomous on Oshana’s account so long as she does not actually give up control (2007: 419).

worry – it is not clear where the ‘relational’ enters my account at all. It seems that reflective endorsement is a process that, theoretically, an agent can undertake on her own. Diachronic value cohesion refers, again, only to factors within the agent. And while I might have saved my account by proposing answerability, I explicitly defend narratability instead – because it does not require that we answer to actual or hypothetical others.

This objection, in turn, leads to the final – and most potentially damaging – objection to my position. Recall the beginning of this chapter, when I discussed how any theory of autonomy should be explicit about what it wants to accomplish, on the one hand, and what it is willing to pay for it, on the other. I recognized that, in order to place priority on women’s self-conceptions, I would have to answer concerns about whether or not my theory can live up to the normative commitments of feminism. The charge is that since my account is internalist and subjective, and since it is not obviously relational *at all*, it will fail in the same way that ‘old-style’ procedural theories failed – it will fail to address the problem of internalized oppression.

In response to an Oshana-style critique, I accept that my theory is internalist, subjective, and psychological, but deny that this is a problem. Oshana’s description of autonomy sounds more like a description of *freedom* to me, where freedom is understood as the absence of barriers or constraints.<sup>101</sup> And while we often equate autonomy with a certain type of freedom – freedom of the will – it is certainly not the same as (what we might call) freedom, full stop. Further, requiring that we have *de facto* control over our lives is much too stringent a requirement. As I noted above, a denial of autonomy is a denial of full agency; when we deny that a person is an agent, we open the door for paternalistic intervention – whether by others or by the state. Imagine if, to prevent such intervention, we had to demonstrate *de facto* control over our lives!

---

<sup>24</sup> Following Isaiah Berlin’s definition of negative liberty.

Oshana's discussion of what this control requires cements my concern. She argues that, in order to be autonomous, a person must have "the power to determine how she shall live" (Oshana 2006: 83). A person lacks this power when "coercive impediments to self-government are *merely likely*, or where it is *possible although unlikely* they will be put into effect"; thus, "autonomy necessitates a fairly robust variety of control of a sort that must be effective within a person's social situation against the presence of *counterfactual impediments*" (ibid., emphasis added).<sup>102</sup> Who among us would be autonomous?

At the heart of Oshana's critique is the idea that, without having an externalist element to my theory, I will fail to take social influences seriously enough. Is my account a relational theory of autonomy at all? If a theory must be constitutively relational in order to 'count as' relational, then my theory will fail to meet this mark. But my theory is *causally* relational; the sorts of relationships in which a person finds herself may in fact influence how she reflects, what her values are, and what kind of story she can tell about these values.

I am somewhat puzzled by the claim that causally relational theories, like mine, which are often procedural and often focus on reflective endorsement, are *too* individualistic.<sup>103</sup> (Constitutively) relational theorists charge that (causally) relational theories are not *really*, or not *sufficiently*, social. In fact, on some accounts, (causally) relational (or procedural) theories are not relational theories *at all*, since they do not build social factors into the definition of autonomy itself.<sup>104</sup> But if we grant that we are all social beings, that our desires and values are all (at the

---

<sup>102</sup> Presumably this is why Meyers finds this condition "so strong that it's not realistically possible to satisfy" (Meyers 2008: 204).

<sup>103</sup> I categorize my theory as one that focuses on reflective endorsement even though I propose two other criteria for autonomy, since reflective endorsement is central to the account.

<sup>104</sup> Talking about constitutively relational theories vs. causally relational theories is a bit cumbersome. Since causally relational theories are usually procedural, I will use these two terms interchangeably. Note, though, that accounts like Westlund's don't fit this mold.

very least) socially influenced, that the tools we use for self-construction are given to us by our relationships, that the very terms we use to understand ourselves are the result of socialization, then the process of reflective endorsement is going to be influenced by our social surroundings as well. This is just to admit that we are social beings – which no one (at this point) denies.

In other words, if you think that it's social *all the way down*, then there is a sense in which it doesn't make sense to charge procedural theories with being too individualistic, as long as these theories acknowledge some of these social aspects. We all have an image of what the caricatured version of reflective endorsement requires – an atomistic, isolated individual (man, probably) walking out deep into the woods to be alone and reflect on his existence, like a hyper-exaggerated John Muir or Henry David Thoreau. This is absurd. But to charge procedural theories with individualism is to think that this is what procedural theorists have in mind. Just because reflective endorsement does not actively involve other people does not mean that it is insufficiently social. Our particular means of endorsement and the feelings that spark critical reflection in the first place are all going to be socially influenced, because *we* are socially influenced. To think that procedural theories are too individualistic is to assume the caricatured version of reflective endorsement. But why must we commit to doing so?

Now, at this point it might look like I've backed myself into a corner. Didn't I embrace a procedural theory in the first place because I wanted a primary focus on individuals? Didn't I charge strong substantive theorists with focusing too exclusively on the social, and not enough on the individual? Now I am saying that the reflective endorsement at the center of my account is, in actuality, social, and that my account is, in actuality, relational? It looks like I'm trying to have my autonomy cake and eat it too. Not at all. What I am pointing out is that those who charge procedural theories with being individualistic are themselves, at least implicitly,



committed to an overly individualistic reading of the reflective endorsement at the center of these accounts. My account, on the other hand, is not committed to a caricatured sort of individualism. It recognizes that we are socially *caused* – which is something about which many traditional accounts were not explicit – without embracing the view that we are socially *constituted*, in an ongoing way, by our relationships. It is thus more individualistic than a majority of feminist relational accounts, and it focuses more on the impact that individuals have on their own autonomy than on the determining effects of social forces. But this is far from embracing the hyper-individualistic ideal of a self-made man.

But does admitting that even the process of reflective endorsement is socially influenced allow my theory to address the problem of internalized oppression? In fact, it might do exactly the opposite. If reflective endorsement is socially influenced, then how could it possibly screen out pernicious forms of socialization? Of course, it is always possible that critical reflection could expose to an agent that her desires are the result of internalized oppression. But it is also possible that the desires are *fully internalized*, so that even the most critical reflection will not reveal their problematic nature to the agent. The requirements of diachronic value cohesion and narratability, though, might help. In some cases, deformed desires will clash with the other values that the agent holds dear, and in other cases agents will not be able to tell a consistent story about how they came to have the desire in question.

But in other cases, neither reflective endorsement, diachronic value cohesion, nor narratability will reveal to the agent that her desire is deformed. These will be the cases in which the agent is so immersed in an oppressive society, for example, that she will not be able to detect that her desires are instilled in her by someone other than herself, that they benefit her oppressors but not her, that they are deformed. But in these cases, where oppressive desires are so fully

internalized by the agent that she is incapable of understanding them as anything but her own, what could it possibly mean to say they aren't? In order to make this argument, we would have to think that there is somehow another her, another self, existing underneath or prior to her socialization. As Suzy Killmister argues, "In refusing to accept the perspective of the deeply socialized agent as authoritative, these criticisms implicitly assume that there must be a true, or authentic, pre-social self lurking within" (Killmister 2013: 99). But in many cases, this is precisely the assumption being criticized. In other words, if oppressive preferences or desires are so deeply internalized that critical reflection could not reveal that they are deformed, then they *just are* the agent's autonomous desires, regardless of their original source; for all of our desires have an original source in something external to us. It is only through the process of internalization that they become our own, and this is no different for deformed desires. Because I recognize that we are socially caused, I can recognize that all desires have an original source in something outside ourselves, before we internalize them and make them our own. Once this process is complete, however – once we have internalized the desires to such an extent that they survive the reflective endorsement, diachronic value cohesion, and narratability criteria – they are authentically ours, and thus, on my account, autonomous.

## BIBLIOGRAPHY

- Abbey, Ruth, ed. *Feminist Interpretations of John Rawls*. University Park, Pennsylvania: Pennsylvania University Press, 2013.
- . *The Return of Feminist Liberalism*. Montreal and Kingston: McGill-Queen's University Press, 2011.
- Alcoff, Linda Martin. "Discourses of Sexual Violence in a Global Framework." *Philosophical Topics* 37, no. 2 (2009): 123-39.
- . *Visible Identities: Race, Gender, and the Self*. New York: Oxford University Press, 2006.
- Anderson, Elizabeth. "Toward a Non-Ideal, Relational Methodology for Political Philosophy: Comments on Schwartzman's *Challenging Liberalism*." *Hypatia* 24, no. 4 (2009): 130-45.
- Anderson, Joel. "Regimes of Autonomy." *Ethical Theory and Moral Practice* 17 (2014): 355-68.
- Arneson, Richard. "Liberal Neutrality on the Good: An Autopsy." In *Perfectionism and Neutrality: Essays in Liberal Theory*, edited by Steven Wall and George Klosko. Lanham, Maryland: Rowman and Littlefield Publishing Group, 2003.
- Babbitt, Susan. "Feminism and Objective Interests: The Role of Transformation Experiences in Rational Deliberation." In *Feminist Epistemologies*, edited by Linda Alcoff and Elizabeth Potter, 245-65. New York: Routledge, 1993.
- Baehr, Amy R., ed. *Varieties of Feminist Liberalism*. Lanham, Maryland: Rowman and Littlefield Publishers, Inc., 2004.
- Baier, Annette. "Cartesian Persons." *Philosophia* 10, no. 3 (1981): 169-88.
- Barclay, Linda. "Autonomy and the Social Self." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 52-71. New York: Oxford University Press, 2000.
- Benhabib, Seyla. *Situating the Self*. New York: Routledge, 1992.
- Benson, Paul. "Feminist Commitments and Relational Autonomy." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 87-113. New York: Oxford University Press, 2014.
- . "Feminist Intuitions and the Normative Substance of Autonomy." In *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy*, edited by James Stacey Taylor, 124-42. Cambridge: Cambridge University Press, 2005.

- . "Taking Ownership: Authority and Voice in Autonomous Agency." In *Autonomy and the Challenges to Liberalism: New Essays*, edited by John Christman and Joel Anderson, 101-26. Cambridge: Cambridge University Press, 2005.
- . "Autonomy and Oppressive Socialization." *Social Theory and Practice* 17 (1991): 385-408.
- . "Feminist Second Thoughts About Free Agency." *Hypatia* 3 (1990): 47-64.
- Brake, Elizabeth. "Rawls and Feminism: What Should Feminists Make of Liberal Neutrality?". *Journal of Moral Philosophy* 1, no. 3 (2004): 293-309.
- Bratman, Michael E. "Hierarchy, Circularity, and Double Reduction." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton, 65-85. Cambridge, Massachusetts: MIT Press, 2002.
- Brennan, Samantha. "The Liberal Rights of Feminist Liberalism." In *Varieties of Feminist Liberalism*, edited by Amy R. Baehr, 85-102. Lanham, Maryland: Rowman and Littlefield Publishers, Inc., 2004.
- Brison, Susan. "Outliving Oneself." In *Feminist Theory: A Philosophical Anthology*, edited by Ann E. Cudd and Robin O. Andreasen, 365-76. Malden, Massachusetts: Blackwell Publishing, 2005.
- Buss, Sarah. "Autonomous Action: Self-Determination in the Passive Mode." *Ethics* 122, no. 4 (2012): 647-91.
- Card, Claudia. "The L Word and the F Word." *Hypatia* 21, no. 2 (2006): 223-29.
- Charles, Sonya. "How Should Feminist Autonomy Theorists Respond to the Problem of Internalized Oppression?". *Social Theory and Practice* 36, no. 3 (2010): 409-28.
- Christman, John. *The Politics of Persons: Individual Autonomy and Socio-Historical Selves*. Cambridge: Cambridge University Press, 2009.
- . "Autonomy, History, and the Subject of Justice." *Social Theory and Practice* 33, no. 1 (2007): 1-26.
- . "Relational Autonomy, Liberal Individualism, and the Social Constitution of Selves." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 117, no. 1/2 (2004): 143-64.
- , ed. *The Inner Citadel: Essays on Individual Autonomy*. Oxford: Oxford University Press, 1989.
- Code, Lorraine. "The Perversion of Autonomy and the Subjection of Women: Discourses on Social Advocacy at Century's End." In *Relational Autonomy: Feminist Perspectives on Autonomy*,

- Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 181-209. New York: Oxford University Press, 2000.
- . *What Can She Know? Feminist Theory and the Construction of Knowledge*. Ithaca, New York: Cornell University Press, 1991.
- Cudd, Ann E. "Wanting Freedom." *Journal of Social Philosophy* 43, no. 4 (2012): 367-85.
- . *Analyzing Oppression*. New York: Oxford University Press, 2006.
- Dworkin, Andrea and MacKinnon, Catharine eds. *In Harm's Way: The Pornography Civil Rights Hearings*. Cambridge: Harvard University Press, 1997.
- Dworkin, Gerald. "The Concept of Autonomy." In *The Inner Citadel: Essays on Individual Autonomy*, edited by John Christman, 54-62. Brattleboro, Vermont: Echo Point Books & Media, 1989.
- Dworkin, Ronald. "Liberal Community." *California Law Review* 77, no. 3 (1989): 479-504.
- Feinberg, Joel. "Autonomy." In *The Inner Citadel: Essays on Individual Autonomy*, edited by John Christman, 27-53. Brattleboro, Vermont: Echo Point Books & Media, 1989.
- . *Offense to Others: The Moral Limits of the Criminal Law*. New York: Oxford University Press, 1985.
- Frankfurt, Harry G. "Reply to Gary Watson." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton, 160-64. Cambridge, Massachusetts: MIT Press, 2002.
- . *Necessity, Volition, and Love*. Cambridge: Cambridge University Press, 1999.
- . *The Importance of What We Care About: Philosophical Essays*. Cambridge: Cambridge University Press, 1988.
- Friedman, Marilyn. "Relational Autonomy and Independence." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 42-60. New York: Oxford University Press, 2014.
- . "Autonomy, Social Disruption, and Women." In *Feminist Theory: A Philosophical Anthology*, edited by Ann E. Cudd and Robin O. Andreasen, 339-51. Malden, Massachusetts: Blackwell Publishing, 2005.
- . *Autonomy, Gender, Politics*. New York: Oxford University Press, 2003.
- . "Autonomy, Social Disruption, and Women." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 35-51. New York: Oxford University Press, 2000.

- . "Autonomy and Social Relationships: Rethinking the Feminist Critique." In *Feminists Rethink the Self*, edited by Diana Tietjens Meyers, 40-61. Boulder, Colorado: Westview, 1997.
- Garnett, Michael. "Autonomy as Social Independence: Reply to Weimer." *Australasian Journal of Philosophy* 93, no. 1 (2015): 155-59.
- . "The Autonomous Life: A Pure Social View." *Australasian Journal of Philosophy* 92, no. 1 (2013): 143-58.
- . "Taking the Self out of Self-Rule." *Ethical Theory and Moral Practice* 16 (2013): 21-33.
- Gaus, Gerald. *Contemporary Theories of Liberalism*. London: Sage, 2003.
- Govier, Trudy. "Self-Trust, Autonomy, and Self-Esteem." *Hypatia* 8 (1993): 99-120.
- Grimshaw, Jean. "Autonomy and Identity in Feminist Thinking." In *Feminist Theory: A Philosophical Anthology*, edited by Ann E. Cudd and Robin O. Andreasen, 329-38. Malden, Massachusetts: Blackwell Publishing, 2005.
- Hay, Carol. *Kantianism, Liberalism, and Feminism: Resisting Oppression*. New York: Palgrave Macmillan, 2013.
- Heikes, Deborah K. *Rationality and Feminist Philosophy*. London: Continuum International Publishing, 2010.
- Herman, Barbara. "Bootstrapping." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton, 253-74. Cambridge, Massachusetts: MIT Press, 2002.
- Hirschmann, Nancy. "Rawls, Freedom, and Disability: A Feminist Rereading." In *Feminist Interpretations of John Rawls*, edited by Ruth Abbey, 40-56. University Park, Pennsylvania: The Pennsylvania State University Press, 2013.
- Jaggar, Alison. *Feminist Politics and Human Nature*. Totowa, New Jersey: Rowman and Alanheld, 1983.
- Khader, Serene J. "Must Theorising About Adaptive Preferences Deny Women's Agency?". *Journal of Applied Philosophy* 29, no. 4 (2012): 302-17.
- Killmister, Suzy. "Autonomy and the Problem of Socialization." *Social Theory and Practice* 39, no. 1 (2013): 95-119.
- . "Autonomy and False Beliefs." *Philosophical Studies* 164 (2013): 513-31.
- Kymlicka, Will. *Multicultural Odysseys: Navigating the New International Politics of Diversity*. Oxford: Oxford University Press, 2009.
- . *Social and Political Philosophy: A Contemporary Introduction*. New York: Routledge, 2002.

- . *Multicultural Citizenship: A Liberal Theory of Minority Rights*. Oxford: Oxford University Press, 1995.
- Langton, Rae. "Whose Right? Ronald Dworkin, Women, and Pornographers." *Philosophy and Public Affairs* 19, no. 4 (1990): 311-59.
- Levey, Ann. "Liberalism, Adaptive Preferences, and Gender Equality." *Hypatia* 20, no. 4 (2005): 127-43.
- Longino, Helen. "Circles of Reason: Some Feminist Reflections on Reason and Rationality." *Episteme* 2, no. 1 (2005): 79-88.
- Lugones, Maria. "Playfulness, 'World'-Traveling, and Loving Perception." In *Feminist Social Thought*, edited by Diana Tietjens Meyers. New York: Routledge, 1997.
- Mackenzie, Catriona. "Relational Autonomy, Normative Authority and Perfectionism." *Journal of Social Philosophy* 39, no. 4 (2008): 513-33.
- Mackenzie, Catriona, and Stoljar, Natalie eds. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*. New York: Oxford University Press, 2000.
- MacKinnon, Catharine. *Toward a Feminist Theory of the State*. Cambridge, Massachusetts: Harvard University Press, 1989.
- . *Feminism Unmodified: Discourses on Life and Law*. Cambridge, Massachusetts: Harvard University Press, 1987.
- Meyers, Diana Tietjens. "The Feminist Debate over Values in Autonomy Theory." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 114-40. New York: Oxford University Press, 2014.
- . "Feminist Perspectives on the Self." In *Stanford Encyclopedia of Philosophy*, 1-81, 2010.
- . "Review Of: Personal Autonomy in Society by Marina Oshana." *Hypatia* 23, no. 2 (2008): 202-06.
- . "Decentralizing Autonomy: Five Faces of Selfhood." In *Autonomy and the Challenges to Liberalism: New Essays*, edited by John Christman and Joel Anderson, 27-55. Cambridge: Cambridge University Press, 2005.
- . *Being Yourself: Essays on Identity, Action, and Social Life*. Oxford: Rowman and Littlefield Publishers, Inc., 2004.
- . *Gender in the Mirror: Cultural Imagery and Women's Agency*. Oxford: Oxford University Press, 2002.

- . "Intersectional Identity and the Authentic Self? Opposites Attract!". In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 151-80. New York: Oxford University Press, 2000.
- . *Self, Society and Personal Choice*. New York: Columbia University Press, 1989.
- . "Personal Autonomy and the Paradox of Feminine Socialization." *Journal of Philosophy* 84 (1987): 619-28.
- Mills, Charles. "Body Politic, Bodies Impolitic ". *Social Research* 78, no. 2 (2011): 583-606.
- Miriam, Kathy. "Stopping the Traffic in Women: Power, Agency and Abolition in Feminist Debates over Sex-Trafficking." *Journal of Social Philosophy* 36, no. 1 (2005): 1-17.
- Nagl-Docekal, Herta. "The Feminist Critique of Reason Revisited." *Hypatia* 14, no. 1 (1999): 49-76.
- Narayan, Uma. "Minds of Their Own: Choices, Autonomy, Cultural Practices, and Other Women." In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, edited by Louise M. Antony and Charlotte E. Witt. Boulder, Colorado: Westview Press, 2002.
- Nedelsky, Jennifer. *Law's Relations: A Relational Theory of Self, Autonomy, and Law*. Oxford: Oxford University Press, 2011.
- . "Reconceiving Autonomy: Sources, Thoughts and Possibilities." *Yale Journal of Law and Feminism* 1 (1989): 7-36.
- Nussbaum, Martha C. *Sex and Social Justice*. Oxford: Oxford University Press, 1999.
- Oshana, Marina. "Autonomy and the Question of Authenticity." *Social Theory and Practice* 33, no. 3 (2007): 411-29.
- . *Personal Autonomy in Society*. Burlington, Vermont: Ashgate Publishing Company, 2006.
- . "Autonomy and Free Agency." In *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy*, edited by James Stacey Taylor, 183-204. Cambridge: Cambridge University Press, 2005.
- Pateman, Carole. "Self-Ownership and the Property in the Person: Democratization and a Tale of Two Concepts." *The Journal of Political Philosophy* 10, no. 1 (2002): 20-53.
- . *The Sexual Contract*. Cambridge: Polity Press, 1988.
- Patten, Alan. "Liberal Neutrality: A Reinterpretation and Defense." *The Journal of Political Philosophy* 20, no. 3 (2012): 249-72.
- Pride, Mary. *The Way Home: Beyond Feminism, Back to Reality*. Crossway Books, 1985.



- Quong, Jonathan. *Liberalism without Perfection*. Oxford and New York: Oxford University Press, 2011.
- Rawls, John. *Political Liberalism*. Expanded edition ed. New York: Columbia University Press, 2005.
- . "Justice as Fairness: Political Not Metaphysical." In *John Rawls: Collected Papers*, edited by Samuel Freeman. Cambridge, Massachusetts: Harvard University Press, 1999.
- . "Justice as Fairness: Political Not Metaphysical." *Philosophy and Public Affairs* 14, no. 3 (1985): 223-55.
- . *A Theory of Justice*. Cambridge, Massachusetts: Harvard University Press, 1971.
- Raz, Joseph. *The Morality of Freedom*. Oxford: Clarendon Press, 1986.
- Sandel, Michael J. *Liberalism and the Limits of Justice*. New York: Cambridge University Press, 1982.
- Scanlon, T.M. "Reasons and Passions." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton, 165-83. Cambridge, Massachusetts: MIT Press, 2002.
- Scheman, Naomi. *Engenderings*. New York: Routledge, 1993.
- Schwartzman, Lisa. "Feminism, Method, and Rawlsian Abstraction." In *Feminist Interpretations of John Rawls*, edited by Ruth Abbey, 40-56. University Park, Pennsylvania: The Pennsylvania State University Press, 2013.
- . *Challenging Liberalism: Feminism as Political Critique*. University Park, Pennsylvania: The Pennsylvania State University Press, 2006.
- Sperry, Elizabeth. "Dupes of Patriarchy: Feminist Strong Substantive Autonomy's Epistemological Weaknesses." *Hypatia* 28, no. 4 (2013): 887-904.
- Stoljar, Natalie. "Autonomy and Adaptive Preference Formation." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 227-54. New York: Oxford University Press, 2014.
- . "Feminist Perspectives on Autonomy." *Stanford Encyclopedia of Philosophy*, no. Spring 2014 (2014): 1-45.
- . "Informed Consent and Relational Conceptions of Autonomy." *Journal of Medicine and Philosophy* 36 (2011): 375-84.
- . "Autonomy and the Feminist Intuition." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 94-111. New York: Oxford University Press, 2000.

- Superson, Anita. "Deformed Desires and Informed Desire Tests." *Hypatia* 20, no. 4 (2005): 109-26.
- Talisso, Robert. *Pluralism and Liberal Politics*. New York: Routledge, 2012.
- Taylor, Charles. "The Politics of Recognition." In *Multiculturalism: Examining the Politics of Recognition*, edited by Amy Guttmann, 25-74. Princeton, New Jersey: Princeton University Press, 1994.
- Taylor, James Stacey, ed. *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy*. Cambridge: Cambridge University Press, 2005.
- Thomas E. Hill, Jr. "Servility and Self-Respect." In *Dignity, Character, and Self-Respect*, edited by Robin S. Dillon, 76-92. New York: Routledge, 1991.
- Velleman, J. David. "Identification and Identity." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton, 91-123. Cambridge, Massachusetts: MIT Press, 2002.
- Watson, Gary. "Volitional Necessities." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton, 129-60. Cambridge, Massachusetts: MIT Press, 2002.
- . "Free Agency." *The Journal of Philosophy* 72, no. 8 (1975): 205-20.
- Watson, Lori and Hartley, Christie. "Is a Feminist Political Liberalism Possible?". *Journal of Ethics and Social Philosophy* 5, no. 1 (2010): 1-21.
- Watson, Lori. "Toward a Feminist Theory of Justice: Political Liberalism and Feminist Method." *Tulsa Law Review* 46, no. 35 (2010-2011).
- . "Pornography." *Philosophy Compass* 5, no. 7 (2010): 535-50.
- . "Pornography and Public Reason." *Social Theory and Practice* 33, no. 3 (2007): 467-88.
- Weimer, Steven. "Autonomy as Rule by the Self." *Australasian Journal of Philosophy* 92, no. 1 (2014): 159-64.
- Weir, Allison. "Who Are We?: Modern Identities between Taylor and Foucault." *Philosophy & Social Criticism* 35, no. 5 (2009): 533-53.
- Westlund, Andrea C. "Autonomy and Self-Care." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 181-200. New York: Oxford University Press, 2014.
- . "Deference as a Normative Power." *Philosophical Studies* 166 (2013): 455-74.

- . "Autonomy in Relation." In *Out from the Shadows: Analytical Feminist Contributions to Traditional Philosophy*, edited by Sharon L. Crasnow and Anita M. Superson, 59-82. New York: Oxford University Press, 2012.
- . "Rethinking Relational Autonomy." *Hypatia* 24, no. 4 (2009): 26-49.
- . "Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?". *The Philosophical Review* 112, no. 4 (2003): 483-523.
- Wheeler, Dwight Furrow and Mark. "Blunting the Blind Impress: Autonomy, Self-Reflection, and Tracking the Truth." *Social Theory and Practice* 39, no. 3 (2013): 477-500.
- Wolf, Susan. "The True, the Good, and the Lovable: Frankfurt's Avoidance of Objectivity." In *Contours of Agency: Essays on Themes from Harry Frankfurt*, edited by Sarah Buss and Lee Overton, 227-44. Cambridge, Massachusetts: MIT Press, 2002.
- . *Freedom within Reason*. New York: Oxford University Press, 1990.
- . "Sanity and the Metaphysics of Responsibility." In *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, edited by Ferdinand David Schoeman, 46-62. Cambridge: Cambridge University Press, 1987.
- Young, Iris Marion. *Intersecting Voices: Dilemmas of Gender, Political Philosophy, and Politics*. New Jersey: Princeton University Press, 1997.
- . *Justice and the Politics of Difference*. Princeton: Princeton University Press, 1990.
- Zerilli, Linda. "Feminist Critiques of Liberalism." In *The Cambridge Companion to Liberalism*, edited by Steven Wall, 355-80. Cambridge: Cambridge University Press, 2015.
- . "Pluralism and the Problem of Judgment: Farewell to Public Reason." *Political Theory* 40, no. 6 (2012): 6-31.
- . "Toward a Feminist Theory of Judgment." *Signs* 34, no. 2 (2009): 295-317.