

*Limiting Heteronomy: An Account of Autonomy to Deal with Oppression*

Maite Rodríguez Apólito

A thesis submitted for the degree of Doctor in Philosophy

Department of Philosophy

School of Philosophy and Art History

University of Essex

2020

## Table of contents

<i>Limiting Heteronomy: An Account of Autonomy to Deal with Oppression</i> .....	1
Table of contents.....	2
Abstract.....	4
Acknowledgements .....	5
Introduction.....	6
Chapter 1: Procedural Models of Personal Autonomy .....	31
I.    Conceptual Landscape.....	31
II.   ‘Internalist’ Proceduralist Models and Their Limits .....	42
III.  Historical Models .....	49
Chapter 2: Theoretical Resources for a Critique of Historical-Procedural Models ....	65
I.    What Counts as ‘Procedural Independence?’.....	68
a)    (Re) defining ‘intersubjective’ procedural independence: .....	72
b)    Problematizing ‘subjective’ procedural independence (i.e. ‘competence conditions’): 82	
II.   What Kind of Historical Materials are Necessary for Critical Self-assessment? 92	
III.  How Should We Understand Oppression?.....	102
Chapter 3: Limits of Historical Models of Personal Autonomy .....	115
I.    Authenticity Conditions – a Critical Assessment.....	118
a)    ‘Authenticity’ - limits:.....	118
b)    Possible rejoinders: .....	127
II.   Consequences of a Critique of Authenticity over Other Elements of the Model	139
a)    Problematizing Christman’s historical condition: .....	139
b)    Problematizing Christman’s competence conditions:.....	147

Chapter 4: Substantive Accounts of Personal Autonomy .....	153
I. The Problem of “Fully-Internalised” Oppression .....	163
II. Beyond the “Full-Internalisation” Model:.....	185
a) Psychological oppression beyond the “full-internalisation” model: .....	191
b) <i>Non-psychological</i> harms of oppression: .....	201
Chapter 5: A Dual-track Approach to Autonomy.....	208
I. Substantive Track – Oppression as ‘Domination’: .....	218
II. Procedural Track – Oppression as ‘Subjection’ .....	225
III. Conclusion .....	240
Bibliography.....	244

## Abstract

Is it possible to limit heteronomy under oppression through critical self-assessment and self-transformation? I answer by testing available models of autonomy in light of their capacity to deal with the forms of heteronomy which typically characterise oppression. Drawing from Foucault's analysis of power relations, I claim that there are significantly different ways of being oppressed in contemporary Western societies and that we need to account for this difference when answering if self-emancipation under oppression is possible.

First, I look into paradigmatic examples of the two main strategies available in the literature on autonomy: Christman's *procedural* account, and Stoljar's and Oshana's *substantive* accounts. I analyse the strengths of these accounts but conclude that, as they stand, they are ill-suited to problematize forms of (what I call) 'subjection', namely forms of oppression which affect agents' "normal" developments qua subjects of different kinds. Crucially, Christman's model lacks resources to problematize settled characters and values, while Stoljar's and Oshana's models cannot sufficiently account for resistance and transgression in oppressive environments.

To find a way out of the impasse of the substantive-procedural debate, I turn to Foucault's analysis of power. Foucault's resources are useful both to problematize agency and self-relations as effects of social power relations and to distinguish between the different interferences that contemporary theorists would call 'oppressive'. I argue that some forms of oppression qualify as Foucauldian 'domination', where power imbalances are frozen and irreversible through the (limited) margin of freedom available to the individuals living in those conditions. Other forms of oppression, however, can be likened to what Foucault calls 'government of individualisation', where practices of self-clarification and self-transformation can make agents less heteronomous vis-à-vis specific power configurations.

I propose a two-tracked approach to autonomy: a revised procedural account for cases of 'subjection', and a substantive one for cases of 'domination'.

## Acknowledgements

I want to express my deepest gratitude to Fabian Freyenhagen. His dedicated supervision improved my work, and his support and mentoring throughout all the different stages of this PhD encouraged me enormously.

I also want to thank all the members of the Philosophy Department at the University of Essex who read earlier versions of some of the chapters of this thesis and provided constructive feedback during panels or presentations. Thanks to Matt Burch, Lorna Finlayson, Timo Jütten, Irene McMullin, and Ellisif Wasmuth. Thanks also to Béatrice Han-Pile for commenting on my project during my first year in Essex. Special thanks to Amy Allen and Lorna Finlayson for a thought-provoking discussion during my viva examination.

I am grateful to all those who taught me philosophy in Uruguay and Paris, in particular to Gustavo Pereira, who helped me secure funding to pursue this PhD; and to Martine Leibovici and Etienne Tassin, who were generous mentors.

I am thankful to Matías Osta Vélez, who supported me during the planning and early stages of this PhD.

I also feel fortunate for having had colleagues who also became close friends: special thanks to Constantin Mehmel and Joshua Kennedy for being great listeners and committed interlocutors.

My warmest thanks to my family back home: Silvana, Gastón, Sofía, and Gabriel. Throughout the years, they have cared for me and encouraged me more than anyone has. I am also extremely grateful to my family far from home: thanks to Madalina Radu for her strength and kindness; and to Ricardo Samaniego de la Fuente, who has not only been a patient and helpful reader but also the best company I could have hoped for during the last year of this PhD.

## Introduction

“Maybe the target nowadays is not to discover what we are but to refuse what we are. We have to imagine and to build up what we could be to get rid of this kind of political ‘double bind,’ which is the simultaneous individualization and totalization of modern power structures.”  
Foucault, “The Subject and Power”.

Throughout this thesis, I explore the possibility of being autonomous or self-governed under oppressive circumstances.<sup>1</sup> In particular, I consider the suitability and limits of what is referred to in the literature as ‘procedural autonomy’ to assess autonomy in oppressive conditions, i.e. I am interested in the possibility of enhancing autonomy under oppression via procedures of reflection carried out by the agent concerned. These processes would include, for example, identifying and limiting the effects of oppressive environments on oneself (e.g. on one’s values or character). In other words, I wish to interrogate the extent to which different forms of self-clarification could contribute to one’s emancipation vis-à-vis oppressive circumstances.

Admittedly, the research question I propose to investigate here (i.e. one that presupposes that there is a possibility for some form of autonomy *even* when one is oppressed) may seem puzzling for several reasons: First, it could be argued that oppression and self-government are not compatible and, indeed, some might even claim that autonomy *requires* the *absence* of oppression.<sup>2</sup> Moreover, it is hard to imagine how self-government could be possible in circumstances so radically oppressive that the “self” that is needed for self-government is severely harmed. Cases where extreme material deprivation or lack of minimal

---

<sup>1</sup> I use “autonomy” and “self-government” interchangeably throughout this thesis.

<sup>2</sup> See for example Natalie Stoljar, “Autonomy and the Feminist Intuition,” in *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, eds. Catriona Mackenzie and Natalie Stoljar (New York: Oxford University Press, 2000); see also Marina Oshana (ed.), *Personal Autonomy and Social Oppression* (New York: Routledge, 2015). I analyse these positions in detail in Chapter 4.

education have compromised one's reflective capacities or seriously damaged one's self-esteem could be examples of the latter. Furthermore, it may be objected that even in "less radical" oppressive circumstances, "self-government" does not necessarily work in an emancipatory way. The role that different forms of self-management play in neoliberal exploitation is a clear example of how self-management could *reinforce* oppression rather than counter it.<sup>3</sup>

These worries are reasonable and will be addressed throughout my thesis.<sup>4</sup> For now, let me briefly anticipate that I plan to draw an important distinction, inspired by the philosophy of Michel Foucault, between different oppressive circumstances – namely, I distinguish between (what I call) 'domination' and (what I call) 'subjection'. (Both terms appear in Foucault's analysis of power relations but, as it will become clearer below, I do not exactly follow Foucault's usage of these terms. Leaving labels aside, the distinction itself is Foucauldian and I justify it by drawing from Foucauldian resources.) This distinction allows me to tackle different oppressive circumstances through different strategies.

Distinguishing between different oppressive circumstances makes sense in light of the aims I set above. As I suggested, I am interested in identifying how and when procedural models could allow for emancipatory forms of self-clarification in oppressive circumstances. Therefore, a first worry should be to determine whether the reflective operations that are typically put in place by procedural models are possible when under oppression. Furthermore, reflection might be possible, but it might not significantly contribute to one's emancipation from specific forms of oppression. I therefore argue that some forms of oppression (i.e. those that fall under the category of 'domination', explained below) are

---

<sup>3</sup> In order to disambiguate these two senses in which one could govern or direct oneself, I use the term "self-management" to refer to forms of self-direction that do not necessarily happen in light of values that I could endorse. I continue to use the term "self-government", in contrast, as a synonym of "autonomy".

<sup>4</sup> I discuss whether autonomy is incompatible with oppression in Chapters 4 and 5; different cases of "harmed" or "powerless" agency in Chapter 4; forms of self-management in light of unproblematized identities in Chapters 3 and 5.

incompatible with a procedural quest to limit oppressive heteronomy.<sup>5</sup> This incompatibility could be explained for at least two reasons: (i) because the “self” needed for reflective self-government is “crushed” by oppression; (ii) because agents, whether or not crushed, are “powerless” to significantly limit their heteronomy vis-à-vis the oppressive circumstances in place. In these extreme cases, an agent’s heteronomy will be decided ‘substantively’ and not procedurally – i.e. it will be decided independently of that agent’s assessment of her own situation and thanks to explicit normative criteria which establishes that certain circumstances are incompatible with autonomy.

However, I argue that other less-extreme oppressive circumstances (i.e. those that fall under the category of ‘subjection’, explained below) are compatible with the task of emancipatory self-clarification outlined above. The reason for this is that ‘subjection’ can be significantly limited through the “right sort” of critical self-assessment, e.g. by modifying self-relations and challenging one’s settled values and preferences. This is why I defend the idea that a *revised* procedural account of autonomy would be apt to carry out the aims I set for my model in *some* oppressive cases.

To unpack the claims made so far, let me clarify some of the key notions and terms at stake, namely ‘oppression’ and ‘self-government’, and guide the reader through the main points of my argument:

As the existing literature on oppression shows, defining this notion is not easy and there is not consensus on the best way to capture the *distinctive* features of oppression, i.e. those that may make oppression a *specific* wrong. One of the most significant disagreements surrounds the possibility of providing a general theory of oppression. That is, theorists discuss whether a comprehensive definition of the phenomenon, i.e. one which encompasses

---

<sup>5</sup> I use ‘oppressive heteronomy’ to refer to those heteronomous interferences or influences which derive from oppressive circumstances.



*all* forms of oppression (e.g. material, political, psychological, cultural), is possible or even useful when it comes to providing “remedies” to oppression.<sup>6</sup>

Without entering into this discussion, there is a minimal agreement when it comes to identifying paradigmatic cases of oppression. Among others, slavery, colonialism, material exploitation, marginalisation, racism, sexism, and other forms of sexual discrimination typically strike oppression theorists as instances of oppression. I will initially work with this “intuitive” understanding of what oppression is to characterise oppression and to make the clarifications and distinctions necessary for my model.

Let me start by introducing some general features of oppression. These features emerge from key points of (significant) consensus in available theories of oppression:

First, oppression is typically thought as a way of significantly “constraining” individuals’ lives through structural “barriers”. For example, those oppressed may be thought as having a severely restricted margin of action. The latter includes cases where options may be available but none of them would significantly change individuals’ situations or would only do so at extremely high costs for individuals. Marilyn Frye, for instance, uses the term “immobilization” to define oppression. According to Frye, oppressed people’s lives are caught up within forces and barriers which “*restrict or penalize* motion in any direction”.<sup>7</sup> For instance, if an agent has to “choose” between inhabiting a demeaning social stereotype *or* being socially invisible, she could be thought as being “immobilized” in the aforementioned sense.

---

<sup>6</sup> See Ann E. Cudd, *Analysing Oppression* (Oxford: Oxford University Press, 2006); and Tamara L. Zutlevics, “Towards a Theory of Oppression,” *Ratio* 15, no. 1 (March 2002) for examples of general theories of oppression; see Iris M. Young, *Justice and the Politics of Difference* (Princeton: Princeton University Press, 1990) for a study which distinguishes amongst different “faces” of oppression; see Nancy Fraser, *Justice Interruptus: Critical Reflections on the ‘Postsocialist’ Condition* (New York: Routledge, 1997), 199 for an argument on how different forms of oppression require different “remedies.”

<sup>7</sup> Marilyn Frye, “Oppression,” in *Gender Basics: Feminist Perspectives on Women and Men*, ed. Anne Minas (Belmont: Wadsworth Publishing, 2000), 12; emphases added.

Second, as it emerges from the above, calling a situation “oppressive” entails a *negative* assessment of the constraints in place. Living a social life that is not constrained *at all*, is likely incompatible with social life. However, oppression is not just any constraint but an *unjust* constraint or, as some authors put it, a way of *unduly* limiting some individuals’ options or margin for action.<sup>8</sup>

Third, oppression affects individuals qua “members” of a group, class, or kind, where this “membership” is, typically, something attributed to the individuals, whether or not it is endorsed by them. Harms that fall upon isolated individuals (even those that may affect their life in significant ways) may be deemed unjust for many reasons but will not be typically deemed ‘oppressive’ if they are not instances of group oppression. For example, Zutlevics considers the case of someone failing to get a job because, during the job interview, they were wearing a shirt of a shade of blue detested by the CEO. We would certainly call such a situation unjust (e.g. because the candidate’s merits were not fairly assessed) but we would not necessarily see it as an instance of oppression.<sup>9</sup> Compare how the situation changes when we imagine that the candidate was not given the job because of their race, gender, or sexual orientation.

Additionally, theorists typically deem some forms of violence and domination as “oppressive” regardless of how representative they are of group oppression. For example, even if someone denied that every case of sex-trafficking is straightforwardly an instance of class oppression, gender oppression or race oppression, they would have to accept that sex-trafficking strikes those made vulnerable by poverty, gender oppression and/or racial oppression disproportionately more than those in privileged positions, and that class, gender, and racial oppression play a decisive role in the emergence and maintenance of phenomena

---

<sup>8</sup> See for example Cudd, *Analysing Oppression*, 4; Zutlevics, “Towards a Theory of Oppression,” 83; Frye, “Oppression,” 13.

<sup>9</sup> Zutlevics, “Towards a Theory of Oppression,” 84.

like sex-trafficking. The key factor here is, I believe, not how common it is for members of a group to be subjected to specific forms of violence or domination, but whether violence or domination derives from (or is facilitated by) oppression-related vulnerabilities.

Fourth, not *every* social constraint on individuals qua members of a group or class is oppressive: in addition to the proviso that constraints need to be illegitimate (noted earlier), constraints that fall upon group members will not be typically deemed oppressive if they provide a net benefit to the group (and its members) concerned. Rather, a group's oppression typically benefits *other groups* at the serious expense of the oppressed. Admittedly, the accuracy of this claim depends a lot on what one means by "benefit" – let me outline my position on this matter: I problematize the dyadic picture of oppression presupposed by the 'benefited vs. harmed' schema in Chapters 2 and 3, and I at points question the assumption that oppression *always* "benefits" the oppressors (as individuals). For example, some might argue that not all individual men benefit from patriarchy in the same way and, even, that men as a collective might be better off without patriarchal values.<sup>10</sup> I consider some of these objections in Chapter 4. For now, let me just anticipate that my model allows for self-assessment and self-clarification in light of social norms and roles that may serve different functions within an oppressive structure. For example, it is possible to question one's character in light of the fact that I happen to occupy a position of privilege in an oppressive structure whether I intend it or not.

Still, what generally holds true is that oppression puts the oppressed (individually and collectively) at important disadvantages in relation to those who are not oppressed in the same way. To put it more bluntly: oppression in its many forms is – as Sandra Bartky puts it when discussing psychological oppression – a way of "fixing disadvantaged persons in their

---

<sup>10</sup> See for example Ania Srinivasan, "Does Anyone Have the Right to Sex?," *London Review of Books*, March 22, 2018. When analysing the case of Elliot Roger (the 'incel' responsible for the Isla Vista killing in California in 2014), Srinivasan discusses how men who do not fit patriarchal (beauty or personality) standards should in fact see feminism as an ally.

disadvantage” or a way to “maintain dominance”.<sup>11</sup> Think for example of the most traditional analysis of waged labour in capitalist societies and, more specifically, of the claim that waged labour is exploitative (and, therefore, oppressive). One common move to criticize this form of oppression consists in pointing to the fact that a capitalist class benefits *at the expense of* workers who have very limited options.<sup>12</sup>

This initial characterisation of oppression is sufficient to start a discussion on whether available models of autonomy provide sufficient resources to deal with oppressive heteronomy, which includes both being able to *identify* oppressive interferences and to give agents an opportunity to *limit* them. One of my proposals is that procedural models as they stand are typically suitable to deal with general (i.e. non-oppressive) heteronomous interventions (e.g. isolated cases of manipulation or coercion) and possibly to identify instances of some forms of extreme oppression (e.g. extreme material deprivation which denies minimal education). However, as they stand, they are ill-suited to deal with oppressive heteronomous interferences that affect one’s general development (e.g. racism or sexism). Let me unpack and justify these claims by guiding the reader through the key steps of my argument:

Firstly, as I mentioned above, I focus on assessing the possibilities and limits of ‘procedural autonomy’. Defining procedural models in detail is one of the aims of my first chapter, but for now let me (broadly) introduce this position as one according to which self-government involves a test which could be potentially satisfied by any preference, desire or

---

<sup>11</sup> Sandra L. Bartky, “On Psychological Oppression” in *Femininity and Domination* (New York: Routledge, 1990), 27.

<sup>12</sup> These “benefits” should be understood broadly, e.g., they are not necessarily economic. For example, Young mentions that, during exploitation, the “energies of the have-nots are continuously expended to maintain and augment the power, status, and wealth of the haves” (Young, *Justice and the Politics of Difference*, 50). Young includes the “transfer of nurturing and sexual energies [from women] to men” as a form of exploitation. (Ibid.) Sandra Bartky also makes use of this idea of illegitimate “transfer” when defining psychological oppression (or ‘psychic alienation’) as an illegitimate “usurpation” of a productive activity that is essential to human nature. According to Bartky, when one is psychologically oppressed, one is “alienated in the production of one’s own person.” (Bartky, “On Psychological Oppression,” 32)

character trait regardless of its content – i.e. procedural tests are *content-neutral*. This means, for example, that we could profoundly disagree with the decisions of autonomous agents or disapprove of their characters. As Gerald Dworkin notes, autonomous agents can be “tyrants or slaves”, “individualists or champions of fraternity”.<sup>13</sup>

And how does this test work? According to this picture, self-government requires that an agent *could* (when necessary) “stand back” from her values or preferences, analyse them critically, and possibly endorse them, “appropriate” them or, at the minimum, not reject them.<sup>14</sup> Consider, for example, cases of what we informally call “pulling oneself together” – i.e. being in (or gaining) control of oneself. First, this task typically involves a form of self-scrutiny, i.e. an effort to gain some clarity on which desires, preferences, and values have a grip on us.<sup>15</sup> Second, there is also self-assessment involved: self-governed agents are not merely expected to be able to describe which desires or preferences have power on them, but they also must typically *approve* of the latter.

Within procedural accounts, the tasks sketched above are generally secured through two conditions: First, a ‘procedural independence’ condition, which grants that agents are *minimally rational* (i.e. they are capable of “good” reflection) and ensures that their reflection is not being illegitimately *interfered* with (e.g. they are not being manipulated).<sup>16</sup> Second, procedural models include an ‘authenticity’ criterion: to count as self-governing, self-direction should happen in light of preferences, desires or emotional tendencies that are

---

<sup>13</sup> Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press, 1988), 29. I analyse ‘content-neutrality’ in more detail in Chapter 1, and problematize it in Chapters 2 and 3.

<sup>14</sup> The qualification with “*could*” is meant to capture the counterfactual proviso of procedural accounts like Christman’s (explained in detail in Chapter 3). The key idea is the following: if I do manage to direct myself without deep conflict or short-circuits but, *should* I assess the values in light of which I direct myself, *I would* feel alienated from them, then I could perhaps say that I have succeeded in “self-management,” but not in ‘authentic self-government’.

<sup>15</sup> Throughout my thesis, I will frequently use the term “content” (e.g. “contents under analysis”) as a general way to referring to the preferences, desires, values, or character traits which may come under reflective scrutiny during self-assessment.

<sup>16</sup> I reflect on these elements, i.e. ‘procedural independence’ and ‘competence conditions’, in much more detail in Chapters 1 and 3.

“one’s own”, i.e. that agents can (or could) *endorse*. Moreover, this endorsement should not happen “carelessly” but should reflect what *really* matters to agents. Crucially, endorsement is not aimed at determining whether a certain desire or preference is *generally* acceptable or “objectively” good (e.g. good for all rational individuals). Rather, a procedural test is meant to secure *individual* endorsement: it invites agents to reflect on whether a value or preference is acceptable given the kind of person one is and wants to be, given the long-standing commitments that one holds, etc. When one cannot (or could not) do the latter in a *deep* sense (i.e. when one is faced not merely with the “normal” ambivalences of everyday life but with an overwhelming sense that a concrete disposition is not in line with “who one is”) then one is alienated from the content in question.

This story of careful self-scrutiny and self-assessment of one’s “inner world” could appear straightforward. However, as many have argued, preferences can be adaptive, desires can be “deformed”, and one’s emotional life more generally is in many ways shaped by one’s context.<sup>17</sup> Indeed, “selves” are (more frequently than one would want to admit) not transparent to themselves.<sup>18</sup> If we cannot “trust” the results of our self-assessments, if things like our preferences, affective reactions, and even the general value frameworks that take a leading role in self-assessment should not be taken at face value, then the picture of authentic self-government described above might appear futile.

Does this mean that models of autonomy should try and “dig deeper”? Should they strive to find contents “untouched” by our contexts and, therefore, *definitively* “ours”? It is

---

<sup>17</sup> See for example, Natalie Stoljar, “Autonomy and Adaptive Preference Formation,” in *Autonomy, Oppression, and Gender*, ed. Andrea Veltman and Mark Piper (Oxford: Oxford University Press, 2014); Ann E Cudd, “Adaptations to Oppression: Preference, Autonomy, and Resistance,” in: *Personal Autonomy and Social Oppression*, ed. Marina Oshana (New York: Routledge, 2014); see Eva Illouz, *Consuming the Romantic Utopia: Love and the Cultural Contradictions of Capitalism*. (Berkeley: University of California Press, 1997) for an argument on how capitalism shapes our emotional lives more generally.

<sup>18</sup> As it will become clear throughout my thesis, talk of “selves”, and even of “true or authentic selves” should not be understood as making any essentialist claims about those “selves”. Indeed, neither me nor the authors that I discuss throughout my thesis (unless stated otherwise) take “selves” to be, for example, metaphysical, unchangeable, entities that exist prior to socialisation.

hard to argue that *any* autonomy theorist has ever seriously defended such a view. Indeed, from the fact that we may be *problematically* influenced by our contexts in many senses, it does not follow that we need to commit only to those preferences, desires or emotional tendencies that were not influenced by our contexts *at all* and which are, so to speak, *originally* ours.

As autonomy theorists rightfully acknowledge, asking for radical originality would leave agents almost (or perhaps even completely) empty-handed: autonomy needs to account for the *fact* that we are *social* beings which means that most of the materials that we could possibly work with to make ourselves more autonomous are made available socially and are socially defined. Moreover, the agentic skills necessary for autonomy are acquired in interaction with others: we *need* others to become autonomous. My research builds upon this widely-accepted premise and should not be read as a call for “radical independence”. Indeed, as it will become clear in the chapters that follow, I suggest that countering oppressive heteronomy is also possible partly thanks to the critical insights on myself and my circumstances gained through the perspectives and circumstances of *others*.

If authentic self-government is not about endorsing only self-made values, what is it about then? Within procedural accounts, ‘authenticity’ is meant to ensure, not radical originality, but the *coherence* that seems to be necessary to be oneself, and to be seen *as* oneself by others.<sup>19</sup> Self-governed agents should generally avoid having conflicting preferences or values at a given time and, even, should be capable of making sense of the (possibly different) values held at different points in their personal histories thanks to the more general organising values presupposed by their histories.

In other words, according to procedural accounts, what is “alien” tends to be that which I cannot integrate into a general (endorsed) value framework or what “short-circuits”

---

<sup>19</sup> Marina Oshana makes the point that the authenticity conditions in some procedural accounts *do* presuppose the possibility of some form of originality. This would be the case, according to Oshana, in Christman’s and Meyers’ theories. (Marina Oshana, “Autonomy and the Question of Authenticity,” *Social Theory and Practice* 33, no. 3 (July 2007), 421. I analyse Christman’s authenticity conditions in detail in Chapter 3.

my motivation for acting “as myself”. In the procedural accounts that I study in Chapters 1 and 3, these general value frameworks are identified thanks to values persistent through time. For example, *sustained* values or *settled* character traits tend to be reliable sources when it comes to identifying ‘authentic’ criteria for self-assessment. Similarly, illegitimate interventions are typically thought in opposition to “normal” developments, namely as interventions that “disrupt” socially-acceptable patterns of socialisation and individualization. In a word, “standard” life developments will not be typically seen as problematic in procedural terms.

When it comes to assessing the sufficiency or insufficiency of the procedural models outlined above, an answer will be, of course, tied to the aims I set earlier. What I mean is that a model of autonomy could be deemed sufficient or not, successful or not, and critical or not, in light of other aims and not (so to speak) “objectively”. In other words, my point is that different ways of defining what counts as “alien” and, indeed, of making sure that *agents themselves* could identify “non-authentic” features of themselves, could be deemed more or less effective depending on what we want to accomplish through a procedural model of autonomy. For example, as I indicate in Chapter 3, John Christman’s model might be perfectly suitable to define what conditions need to be met for an agent to be granted liberal rights and, consequently, to decide when paternalistic interventions are unjustified.<sup>20</sup> However, *if* we want to secure the conditions of possibility for limiting one’s oppression, then Christman’s model is in need of further elaboration. Crucially, I argue, the procedural solution sketched above is dangerous in the context of a discussion on the possibility of limiting oppressive heteronomous interferences which act on one’s general development as a subject. Let me justify this claim below:

---

<sup>20</sup> Other aims pursued through a theory of autonomy could be: identifying when agents are responsible for their actions, identifying which sorts of actions promote fulfilment or self-realization, identifying conditions of legitimate consent (e.g. medical), among others. See Diana T. Meyers, “Feminist Debates Over Values in Autonomy Theory,” in *Autonomy, Oppression, and Gender*, 133-4.



What counts as a “standard” life development will be typically defined in a way that is functional to the social configuration in place. In light of this, when it comes to defining what it would take for an agent to be able to limit oppressive circumstances procedurally, we need to assess the orienting role that “normal” or “standard” developments frequently play in procedural models. Similarly, we need to be cautious about the role assigned to sustained or settled commitments or personal features.

The case of gender oppression provides a good example to clarify my abovementioned point: being socialised as a ‘woman’ in a patriarchal context happens by way of socialization processes which are generally deemed acceptable by the relevant community. That is, an oppressive development in such circumstances would not necessarily involve big disruptions or obviously illegitimate interferences.<sup>21</sup> Moreover, having been socialised as a woman in a patriarchal context surely is a source of many sustained character traits and settled values and preferences which are (as many feminists would agree) *also* part of the problem if one aims at tackling female oppression. Therefore, identifying these sustained and settled features of oneself as ‘authentic’ by virtue of their persistence throughout one’s development would leave us in a tricky position should we want to limit oppression and its effects on ourselves.

Understandably, a committed proceduralist might feel inclined to play the endorsement card again: I could make these developments themselves relevant material for the subjective endorsement (or lack thereof) which is necessary for procedural autonomy. That is, a proceduralist could concede that normal developments might be problematic in many ways (especially if one takes oppression seriously), but whether or not they enable autonomy is something that agents *themselves* should decide. Leading a self-governed life, the

---

<sup>21</sup> See for example Amy Allen’s description of Elizabeth’s case (a girl socialized according to oppressive gender norms in spite of her parents’ good intentions) in Amy Allen, “Recognizing Domination: Recognition and Power in Honneth’s Critical Theory,” *Journal of Power* 3, no. 1 (April 2010). Allen argues that it is possible to receive parental love and gender subordination “in a single stroke.” (Ibid., 26)

argument goes, is not necessarily about living life according to “good” norms or values, but rather about following *self*-given norms or values (i.e. those acceptable according to *the agent’s own* standards). For example, if a woman looks back at her personal development and is able to come to terms with the values she holds and the preferences she has *all things considered* (e.g. even in light of the fact that her preferences adapted to a patriarchal context and upbringing) then she should count as autonomous.

The latter is an alternative defended by *historical* accounts of personal autonomy and, more specifically, developed in Christman’s influential model. According to Christman, the self-assessment necessary for autonomy should happen in light of one’s personal history. For example, a certain disposition might be, by virtue of its content, compatible with everything I stand for but I could *still* feel alienated from this disposition because, say, I learn that the disposition is the result of a certain upbringing that is *not* compatible with my value framework.

I argue that, as promising as historical procedural accounts may be in many respects (analysed in Chapter 1), they do not secure a viewpoint critical enough to ensure that agents have the *opportunity* to assess oppressive heteronomous interferences. The reason for the latter is that an endorsement (or lack thereof) which follows a self-assessment carried out in light of one’s *personal* history alone will fail to provide, for example, enough information on the social meaning or function of one’s personal development. Moreover, challenging what is deemed “personal” in the first place is crucial when it comes to limiting oppressive circumstances.

Consider the case from above: a woman who endorses her values in light of her personal history. My suggestion is that, if her endorsement occurs without sufficient awareness of the *social* conditions of possibility of her personal or psychological history, this endorsement should not be considered enough for the purposes of limiting gender

oppression. Unpacking and arguing in favour of the latter claim will be key throughout my thesis but let me briefly anticipate some of the reasons that support this claim:

Take again the brief example proposed above, what is missing in this agent's assessment? This agent could endorse features of herself in light of her personal history without acknowledging, for instance, that hers is not an isolated biography or that her upbringing "simply" followed social guidelines for raising daughters. And why should this matter? I argue that this matters greatly for the purposes of limiting oppressive interferences: placing one's personal history within a wider social history is crucial to unveil the social roles that structure, enable, and limit the stable "personal" features on which historical procedural accounts focus.

Crucially, having an opportunity to interrogate stable values, sustained character traits, and other persistent features of oneself matters to secure the conditions of possibility for experiencing *alienation* from the latter in oppressive contexts which shape one's general development. Indeed, I argue that experiencing alienation (which includes: identifying discomforts or sufferings *as* experiences of alienation) from one's organizing values in oppressive circumstances which shape one's development is made difficult by the notion of 'authenticity' as it stands in procedural accounts. If experiencing alienation is not sufficiently made possible, then the results of a historical procedural test cannot be considered authoritative for determining an agent's autonomy vis-à-vis oppressive heteronomous interferences.

At this point, someone could rightfully wonder if it is sensible to hold on to a procedural project in the context of this discussion. After having discussed the different ways in which procedural accounts of autonomy could be deemed problematic or insufficient, it seems only reasonable to consider alternative approaches to autonomy. As I said earlier, my thesis also considers *substantive* approaches to autonomy, namely those which decide autonomy ultimately by reference to explicit normative commitments which define, for

example, what the *world* should look like for autonomy to be possible. Crucially, substantive conditions for autonomy hold *regardless* of whether the agent whose autonomy is being assessed agrees with them. In Chapter 4, I consider Marina Oshana's, Natalie Stoljar's, and Paul Benson's accounts as examples of substantive strategies which limit either: the kinds of social relations compatible with autonomy (Oshana), the content of the social norms or preferences compatible with autonomy (Stoljar), or the "quality" of the self-relations which are compatible with autonomy (Benson).

Substantive theorists of autonomy have raised critiques against procedural accounts motivated by concerns similar to the ones analysed earlier: agents' *psychologies* are much more affected and conditioned by oppressive *social* elements than standard procedural models can account for. For example, I assess Stoljar's position on the effects of oppressive stereotypes on autonomy, namely that one cannot have autonomous preferences in pernicious normative environments.

While I am sympathetic to the problems raised by substantive theorists, I am not in full agreement with them on a crucial point: as I said before, I do not believe that oppression requires us to adopt a substantive position in *all* cases. I accept that in severe cases of oppression no 'independence of mind' could be enough to secure autonomy insofar as this independence could not, when oppression is of a particularly severe kind, guarantee the necessary control over one's life that is crucial for meaningful self-government. (Oshana's theory will provide valuable insights to identify the kind of control over one's life necessary for self-government.) Nonetheless, my point is also that in other oppressive contexts (e.g. those that Stoljar sees as problematic) the appropriate critical self-assessment *could* make a *qualitative* difference vis-à-vis limiting one's oppression. Being subjected to racist or sexist stereotypes would be a case that falls within the latter category. Additionally, I say that critical self-assessment can progressively lead to the more "external", social, and relational change that substantive theorists (rightfully) perceive as necessary to fully tackle oppression.

Crucially, there is another *ethical* reason why I believe that we should prefer a procedural approach in many cases of oppression and limit substantive accounts for the most radical cases of oppression. I believe that once we have established criteria to make individual reflection less vulnerable to those social or relational influences that could make agents self-assessments “untrustworthy”, agents’ self-assessments and experiences should be given special *authority*. This point is relevant when defining autonomy for many contexts (e.g. for medical contexts) but it is particularly pressing when reflecting on the possibility of limiting heteronomy when *oppressed*. As I argue throughout my thesis, suggesting alternatives to deal with oppression in a way that disregards the voice and experiences of the oppressed, may end up “adding insult to injury” and even reinforcing oppression.<sup>22</sup>

To summarize my points so far, my aim is to give *both* procedural and substantive accounts their due. I concede that autonomy, as a form of reflective self-clarification, is neither always possible nor a remedy to all forms of oppression. Indeed, claiming that oppression always can (or should) be significantly reversed through mechanisms carried out by those who are oppressed, might oversimplify what oppression is. Moreover, in many cases, arguing that oppression can be reverted by subjective means could end up overburdening those oppressed. Some forms of oppression need external or infrastructural change to happen before self-government is possible in a meaningful way (e.g. for it not to merely amount to self-management).

The latter, however, does not exclude that procedural processes of self-clarification could have a crucial role when countering *some* forms of oppression that are particularly common in contemporary liberal societies. Nonetheless, I also argue that taking the danger of self-management seriously requires that we go *beyond* available procedural models.

---

<sup>22</sup> I am not suggesting that all substantive theorists end up reinforcing oppression or that they do so in the same ways. I will discuss different substantive models in Chapter 4. My point for now is that the general substantive strategy could have dangerous setbacks when thinking about autonomy in contexts of oppression.

Crucially, this is so even if we defend the view (as I do in this thesis) that procedural accounts can have an emancipatory potential. What I mean is that even in relation to oppression that could be (partially) reversed by way of procedural autonomy, procedural accounts need important reworking. In particular, procedural accounts need to define more carefully the conditions that are necessary for agents to be able to *critically* assess their situations and to articulate different forms of affective discomfort or suffering as experiences of alienation.

As I said earlier, in order to disentangle the different ways in which oppression works and to decide on the strategies that are most suitable to deal with them, I draw on conceptual tools from Foucault's analysis of power relations. A Foucauldian framework allows us both to (i) distinguish between different forms of oppression, *and* (ii) to work under the assumption that governing oneself is (very frequently) also a way of being governed by others. These features, I believe, make this framework particularly promising to find a way out of the impasse of the substantive-procedural debate.

Regarding point (i), although Foucault rarely uses the term "oppression", he frequently refers to circumstances that contemporary theorists would deem oppressive – e.g. colonialism, racism, the power that men exert over women, institutional abuses of power, among others. Admittedly, these references are not always sufficiently developed and Foucault sometimes overlooks elements necessary to carry out an analysis of oppression in contemporary societies. For example, Bartky notes that Foucault does not deal in detail with the production of "feminine" bodies and subjects and Thomas Holt points out that Foucault overlooks the crucial phenomenon of colonialism when dealing with race and racism.<sup>23</sup> Still, these same authors argue that Foucault's own framework can be used productively to fill in

---

<sup>23</sup> Sandra L. Bartky, "Foucault, Femininity, and the Modernization of Patriarchal Power," in *Femininity and Domination*, 79; Thomas Holt, "Pouvoir, savoir et race: À propos du cours de Michel Foucault 'Il faut défendre la société'," in *Lectures de Michel Foucault, Vol. 1*, ed. Jean-Claude Zancarini (Paris: ENS Editions, 2001). See also Ladelle McWhorter, "Post-Liberation Feminism and Practices of Freedom," *Foucault Studies*, no. 16 (September 2013) for an analysis of the relations between contemporary theories of oppression (Marilyn Frye's in particular) and Foucault's concepts of government and domination.

existing lacunae and that it is possible to account for contemporary forms of gender and race oppression along Foucauldian lines. In a word, there is sufficient overlap between Foucault's areas of concern, Foucauldian scholarship, and contemporary debates in autonomy and oppression, to justify bringing these corpora of works together.

Furthermore, in the context of his analysis of strategic power relations, Foucault makes a three-level distinction between (1) dynamic strategic relations, (2) techniques of government, and (3) states of domination.<sup>24</sup> These three levels could be distinguished along two dimensions: (a) the amount of individual freedom that they presuppose (it is abundant in strategic relations, present in techniques of government, and extremely limited in states of domination); and (b) the extent to which they are reversible or dynamic (this includes notably the extent to which they can be reversed by using the margin of freedom available to the individuals within these relations). Foucault suggests in late interviews that (2) and (3) are games of strategy which have been, in a specific society, stabilized or "frozen" altogether: social power (i.e. what could be here linked to 'techniques of government') is a "strategic relation which has been stabilized through an institution",<sup>25</sup> and 'states of domination' are social power relations which have been "blocked" or "frozen" by an individual or group.<sup>26</sup> Paradigmatic examples of each configuration would be: (1) dynamic strategic games which may arise in everyday contexts when one is trying to influence the conduct of others (e.g. during an interview); (2) more institutionalised and therefore stable social power relations that allow for the government of populations as long as individuals conform to specific patterns (e.g. some forms of psychiatric power and patriarchal power); (3) ossified macro-

---

<sup>24</sup> Michel Foucault, "The Ethics of the Concern of the Self as a Practice of Freedom," in *Ethics, Subjectivity, and Truth, Vol 1 of The Essential Works of Michel Foucault 1954-1984*, ed. Paul Rabinow (New York: The New York Press, 1997), 299.

<sup>25</sup> Michel Foucault, "Sex, Power, and The Politics of Identity," in *Ethics, Subjectivity, and Truth, Vol 1 of The Essential Works of Michel Foucault 1954-1984*, ed. Paul Rabinow (New York: New York Press, 1997), 169.

<sup>26</sup> Foucault, "The Ethics of the Concern of the Self as a Practice of Freedom," 283.

level power relations that are maintained, for example, through direct coercion or legal barriers (e.g. colonialism).

I propose that what we (and contemporary theorists) would typically call ‘oppression’ happens mostly on levels (2) and (3) but suggest that we keep these levels separate for the purposes of analysing what would be required to limit oppression in each case. Keeping these instances of oppression separate has a crucial advantage in the context of this thesis: distinguishing states of domination from forms of government allows us to unveil the different extents to which these configurations benefit from individuals being certain *kinds* of subjects or, to put it differently, from shaping themselves and exercising their agency in specific ways. From now on, as I explained earlier, I will refer to these two different levels as ‘subjection’ (level 2) and ‘domination’ (level 3). Since these terms are extremely loaded within Foucault’s work and Foucauldian literature allow me some terminological clarifications:

As it is well-known, Foucault uses the term *assujettissement* – i.e. ‘subjection’ – to refer to the fact that subjects are the *products* of social power relations. Processes of subject-formation are themselves technologies of power and, according to Foucault, it makes no sense to think of subjects as existing prior to power relations or completely independently of them. In other words, individuals are not “primitive atoms” to which power is applied.<sup>27</sup> As Foucault puts it: “there isn’t power on the one hand and then the people to whom power is applied on the other.”<sup>28</sup>

If we take the claims presented above too pessimistically, we could conclude that individuals are completely determined by power. Indeed, commentators typically resort to Foucault’s early analysis of subjection (e.g. those carried out when studying the disciplinary techniques of the seventeenth and early eighteenth centuries) to make the point that,

---

<sup>27</sup> Michel Foucault, *Society Must be Defended: Lecture at the Collège de France 1975-6*, trans. David Macey, (New York: Picador, 2003), 29.

<sup>28</sup> Michel Foucault, “There Can’t be Societies without Uprisings,” in *Foucault and the Making of Subjects*, ed. Laura Cremonesi et al. (Lanham: Rowman and Littlefield International, 2016), 38.



according to Foucault, modern individuals are nothing or little more than docile bodies.<sup>29</sup> Commentators who wish to avoid the latter interpretation frequently resort to statements made by Foucault later in his career where he claims that, even in his early writings on disciplinary power, processes of subject-formation always implied a dimension of *self*-constitution. Moreover, Foucault insisted on the fact that he never meant to reduce subjects neither to objects (i.e. to bodies exclusively) nor to completely passive receptors of power.<sup>30</sup> Indeed, even when outlining his model of disciplinary power, Foucault remarks that individuals are not “inert matter” merely “struck” by power.<sup>31</sup>

In this thesis, I focus mostly on Foucault’s (later) reflections on government. As I explain in Chapter 2, in the context of his analysis of government, Foucault problematizes more explicitly the role that self-relations, accepted truths on oneself, and different forms of self-scrutiny play in individuals’ subjection. Foucault famously claims: “There are two meanings of the word “subject”: subject to someone else by control and dependence, and tied to his own identity by a conscience or self-knowledge”.<sup>32</sup> One key issue at this point is that the ways in which individuals self-relate and the use that individuals make of the freedom available to them *ties* them to *others*. That is, individuals constitute *themselves* and self-manage in a way that makes them governable in light of social “truths” and goals. One of the main

---

<sup>29</sup> For example, Foucault claims in 1973: “disciplinary power, and this is no doubt its fundamental property, fabricates subjected body; it pins the subject function exactly to the body. [...] It is individualizing [in] that the individual is nothing other than the subjected body”. (Michel Foucault, *Psychiatric Power: Lectures at the College de France 1973-74*, trans. Graham Burchell (Basingstoke: Palgrave Macmillan, 2006), 55. In *Discipline and Punish* Foucault refers to the “internal” dimension of modern individuals (i.e. to their “souls”) as an effect of the training of bodies. (Michel Foucault, *Discipline and Punish: The Birth of the Prison* (New York: Vintage Books, 1995), 24.

<sup>30</sup> See for example: “If it is indeed true that the constitution of the mad subject may be considered the consequence of a system of coercion – this is the passive subject – you know very well that the mad subject is not an unfree subject, and that the mentally ill person constitutes himself [*se constitue*] as a mad subject precisely in relation to and over against the one who declares him mad.” (Foucault, “The Ethics of the Concern of the Self as a Practice of Freedom,” 291; translation modified). See also Laura Cremonesi, ed., *Foucault and the Making of Subjects*. (Lanham: Rowman and Littlefield International, 2016) for a collection of essays which analyses ‘subjection’ vis-à-vis ‘subjectivation’ (active processes of self-constitution).

<sup>31</sup> Foucault, *Society Must be Defended*, 29.

<sup>32</sup> Michel Foucault, “The Subject and Power,” in *Power Vol. 3 of Essential Works of Michel Foucault 1954-1984*, ed. James D. Faubion (New York: The New York Press, 2000), 331.

aims of this thesis is to explore how it might be possible to limit one's own subjection and therefore the government by others enabled by subjection.

Admittedly, for some readers the possibility of limiting subjection through subjective means might appear impossible within a Foucauldian framework. A research project that deals with the possibility of autonomy via a Foucauldian framework needs to deal with the issue of whether subjects could ever “de-subject” themselves. After all, within a Foucauldian framework, any self-clarification procedure (like the ones described above) should be problematized: according to Foucault, different ‘techniques of the self’ work *within* specific governmental configurations – i.e. they make sense and play a role in light of the broader aim of governing individuals and populations according to, e.g., social or economic aims. Therefore, by governing oneself – an argument could go – one in fact *increases* rather than *limits* one's submission to a power configuration. A classical illustration of the latter is the relationship between different forms of “responsible self-management” and neoliberal governmentality described in the beginning of this introduction. Foucault (and others after him) could argue that the many ways in which individuals are made responsible for their health, their retirement, and are expected to adapt to fast-moving working conditions make individuals functional to the systemic logic of neoliberalism. As Lois McNay summarizes: “If individual autonomy is not the opposite of or limit to neoliberal governance, but rather lies at the heart of disciplinary control through responsible self-management, what are the possible grounds upon which political resistance can be based?”<sup>33</sup>

Authors like Amy Allen have argued that Foucault's philosophy can accommodate ideals of autonomy and resistance.<sup>34</sup> According to Allen, two “twin notions of autonomy”

---

<sup>33</sup> Lois McNay, “Self as Enterprise: Dilemmas of Control and Resistance in Foucault's ‘The Birth of Biopolitics,’” *Theory, Culture, and Society* 26, no. 6, (2009): 56.

<sup>34</sup> Allen also argues that these ideals can be accommodated “consistently” by Foucault's work, insofar as Allen defends the idea that Foucault's work *always* offered room for ideas like ‘subjectivity’ and ‘autonomy’, i.e. this possibility is not due to a radical shift in Foucault's philosophy. Amy Allen, *The politics of Our Selves: Power, Autonomy, and Gender in Contemporary Critical Theory* (New York: Columbia University Press, 2008).

can be extracted from Foucault's work: autonomy can be thought both as a capacity for critical reflection and as a capacity for self-transformation.<sup>35</sup> This means that gaining a "critical insight" on power (i.e. on power relations and on the self-relations associated with specific power configuration) is possible *and* that agents are capable of self-transformation in light of these critical insights.<sup>36</sup>

I engage with Foucault's notion of the 'critical ethos' and with the possibility and potential of self-transformation in Chapter 2. For now, let me just briefly say that 'critique' in Foucault implies a special 'historical ontology of ourselves', which allows us to unveil the historical *contingency* of certain forms of being that we could have otherwise thought necessary. Furthermore, Foucault connects this critical ethos or attitude with the task of *limiting* government by others. The possibility of limiting government should be interpreted with caution: a life completely unaffected by power is impossible within a Foucauldian schema, it is not possible to choose not to be governed *at all*.<sup>37</sup> However, Foucault grants that there is room for forms of local resistance. Indeed, unless we are dealing with outright domination, it is possible to make oneself *less* governed (i.e. more autonomous) vis-à-vis configurations that one finds particularly problematic or undesirable (e.g. a particular power relation that makes one functional to social ends or structures that one rejects).<sup>38</sup>

Regarding the claim that, according to Foucault, subjects are capable of self-transformation, I provide a brief explanation for now: as I said before, subjection is not

---

<sup>35</sup> Ibid., 2. Indeed, the term "autonomy" is explicitly used by Foucault in the context of his analysis of the "modern ethos" or the "critical ethos." (Michel Foucault, "What is Enlightenment?," in *Ethics, Subjectivity, and Truth, Vol 1 of The Essential Works of Michel Foucault 1954-1984*, ed. Paul Rabinow (New York: New York Press, 1997), 313). Moreover, Foucault appeals to a form of "self-government" (*maîtrise de soi*) which does not seem reducible to a form of self-management in the context of his analysis of free self-constitution and the care of the self. (Foucault, "The Ethics of the Concern of the Self as a Practice of Freedom," 301). I come back to these two contexts in Chapter 2.

<sup>36</sup> Allen, *The politics of Our Selves*, 46.

<sup>37</sup> Michel Foucault, "What is Critique?," in *The Politics of Truth*, ed. Sylvère Lotringer (New York: Semiotext(e), 1997), 28.

<sup>38</sup> In fact, Foucault's claim that critique is « l'art de n'être pas *tellement* gouverné » translates better as "the art of not being governed *quite so much*" (which would imply the possibility of being governed *less*) than as "the art of not being governed *like that*" (which would simply imply a choice between different forms of being governed). (Ibid, 29; emphases added)

something that merely “happens” to individuals. Being “subjected” implies that one is capable of a number of actions on oneself and that one has made use of the margin of freedom available in that process in a certain “directed” way. Subjection, therefore, presupposes “skilful” agency and, crucially, entails the possibility of using those skills otherwise – e.g. for ends other than one’s subordination to social control.

Finally, it should be remembered that Foucault repeatedly warns his readers against a simplistic reading of his notion of ‘power’: power is not an “evil substance”,<sup>39</sup> it is not a synonym of outright violence, it is not something that some “hold” while others “lack” completely.<sup>40</sup> As I explain in Chapter 2, “power” should be thought as a *relation*, and is a way of directing the use that agents make of their *freedom*, which is to say that power relations presuppose that acting otherwise is possible.<sup>41</sup> In other words, one can make the case that power relations in general (including those that are more stable and fall within the category of ‘governmental techniques’) “tame” subjectivity.<sup>42</sup> This “taming” is indeed a danger but should not be seen as an irredeemable one: becoming subjects differently, more critically, or even becoming other kinds of subjects, is possible. Indeed, in late texts, Foucault considers the possibility of collectively and individually refusing, challenging, and resisting forms of subjection.<sup>43</sup>

I hope this brief discussion suggests that ‘subjection’ need not be seen under an overly-pessimistic light. Let me warn the reader, however, that on these interpretative matters I rely greatly on the work of Foucault scholars. That is, I discuss and comment on Foucault when relevant and necessary, but I do not engage in detail with discussions internal to Foucault scholarship to convince the reader that thinking about self-government,

---

<sup>39</sup> Indeed, my use of “power” (unless stated otherwise) should be understood as implying a “power relation” and not power as a “substance”.

<sup>40</sup> Foucault, *Society Must be Defended*, 29.

<sup>41</sup> Foucault, “The Subject and Power,” 340.

<sup>42</sup> The use of the word “tame” should not suggest to the reader the existence of a subjectivity *prior* to power.

<sup>43</sup> See *Ibid.*, 336; 342.

emancipation, and resistance in connection to Foucault is viable. As I have explained, I believe it is and that other authors have successfully made this point.<sup>44</sup> Similarly, I leave it to others to discuss in detail whether Foucault's late writings are compatible or incompatible with his earlier ideas. This is not a thesis *about* Foucault but a thesis that tests available models of autonomy in light of the opportunities they provide to deal with oppression, where oppression is conceptualised in a Foucauldian-inspired fashion.

To summarize, what should be retained for now is that my use of 'subjection' does neither entail the complete passivity or object-status of the oppressed (i.e. they are agents) nor a priori forecloses the possibility of developing individual and collective strategies to limit subjection. Namely, when we are faced with what I label 'subjection', there is enough freedom for subjects to modify themselves and to conduct *themselves* in different ways and power relations are (at least partially) *reversible* by subjective means. What should be clear, however, is that critical self-assessment and self-transformation can only happen through materials and subjectivities which are a product of oppression (though not entirely determined by oppression).

Regarding 'domination', as I mentioned earlier, I use this term not as a straightforward synonym of 'power relation' but to denote some extreme forms of oppression where freedom is severely reduced. In these extreme oppressive cases, *external* limits prevent lives from being shaped in light of self-governed decisions or in light of the values that one endorses. In chapter 2, I argue that these extreme cases where power relations are no longer open-ended and asymmetries are no longer significantly reversible could occur both on a macro (social) level and on a micro (interpersonal) level. On the macro-level, ossified power relations (or, more specifically, ossified governmental techniques) derive in 'states of domination' (defined above). On the micro-level, strategic power relations can be

---

<sup>44</sup> As I mentioned earlier, a noteworthy example is Allen, *The Politics of Our Selves*.

ossified when power imbalances are extremely asymmetrical or when one of the parties within the strategic relation resorts to sheer violence. My use of ‘domination’ therefore includes three phenomena considered by Foucault: ‘states of domination’, radical power asymmetries which leave agents incapable of “playing” the games of power, and ‘relationships of violence’.<sup>45</sup> I argue that, when individuals are dominated, forms of self-clarification, no matter how critical, will *not* be sufficient to limit oppression.

Tying the distinction between subjection and domination to my earlier discussion, I propose (in Chapter 5) a two-tracked approach to autonomy: I argue that a (revised) procedural account could help to limit oppressive heteronomy when one is subjected but not when dominated. When domination is the case, autonomy should be decided substantively. All in all, I argue that testing the viability of available models to deal with oppression requires both that we take the dangers of domination and subjection seriously and that we make sure that, where limiting oppression through self-clarification is possible, self-clarification is critical enough, and affectively compelling enough, to do so.

---

<sup>45</sup> Admittedly, not all ‘relationships of violence’ are instances of oppression (e.g. an individual may be a victim of violence which is neither the result of her membership to an oppressed group nor is explained by the vulnerabilities associated with her belonging to an oppressed group). As it will become clearer later, including cases of “isolated” violence on the substantive track of my model is not problematic. Indeed, cases of sheer violence would typically fail to meet the procedural independence conditions set by available procedural models. My domination track, therefore, captures both the obvious cases of manipulation and deception which are generally contemplated by procedural models as they stand and, additionally, it takes into account some extreme forms of structural oppression.

## Chapter 1: Procedural Models of Personal Autonomy

### I. Conceptual Landscape

I start by analysing (so-called) ‘procedural’ strategies to define personal autonomy. According to procedural models, the autonomy of an agent vis-à-vis a disposition, act, or character trait typically involves corroborating whether the disposition, act, or character trait in question has a certain “status” within the agent’s psychology or motivational structure. Namely, it is crucial that these dispositions, acts, or character traits are “one’s own” (i.e. ‘authentic’) in different senses that will be explained below. This (rough) schema allows to capture the overall strategy of procedural accounts but, for the sake of giving the reader a more detailed map of the theoretical landscape, allow me some initial clarifications:

First, the process through which this “corroboration” happens varies greatly depending on the particular procedural model. For example, autonomy “checks” can be understood as the intellectual acts of disembodied selves which occur at a specific point in time; as an affective and cognitive test carried out by socio-historical selves at a specific point in time; as a dialogical test in which I answer for my values to (real or imaginary) others; or as a process carried out by multi-dimensional (e.g. embodied, social, relational, cultural) selves.<sup>46</sup>

Second, *what* is exactly under scrutiny also depends on the particular procedural model: some models focus on the autonomy of agents vis-à-vis more or less stable ‘volitions’ (e.g.

---

<sup>46</sup> See Harry Frankfurt, “Freedom of the Will and the Concept of a Person” in *The Inner Citadel: Essays on Individual Autonomy*, ed. John Christman (New York: Oxford University Press, 1989); John Christman, *The politics of persons* (Cambridge: Cambridge University Press, 2009); Andrea Westlund, “Rethinking Relational Autonomy,” *Hypatia*, 24, no. 4 (2009); Diana T. Meyers, *Being yourself: Essays on Identity, Action, and Social Life* (New York: Rowman & Littlefield Publishers, 2004) for examples of each strategy respectively. Because of the aims of my model (dealing with oppression even when the latter affects individuals’ general developments), I focus mostly on the second strategy. As it will become clearer in what follows, I believe that we cannot deal with oppression without considering the historical conditions of emergence of characters and values.

desires or preferences), while others assess values and, crucially, those settled value-orientations and commitments that “organise” an agent’s life. I will indicate, when discussing different models, what kind of object is considered in each case.

It is worth noting, however, that when procedural theorists disagree about which particular contents should be under analysis to determine an agent’s autonomy, this disagreement does not merely reflect the fact that theorists happen to focus on different “objects”. Rather, these differences also frequently reveal a crucial difference in *scope* between available theories. That is, while some theorists focus on what makes particular *acts* (e.g. particular choices) autonomous, others focus more generally on what makes the *values* or *characters* that orient our lives (and, indeed, what makes these lives themselves) autonomous.

Let me anticipate that I favour a discussion which allows us to question the autonomy of ongoing human lives, of lifestyles, of developing characters, and not exclusively of particular choices. In other words, I am interested in an analysis that could shed light on what is to be autonomous in (what Diana Meyers calls) a ‘programmatically’ way, not just in an ‘episodic’ one. I clarify Meyers’ distinction below:

Meyers illustrates the distinction between ‘programmatically’ and ‘episodic’ autonomy by contrasting the scope and the implications of two different questions: ‘How do I want to live my life?’ vs. ‘What do I want to do now?’<sup>47</sup> These questions capture, respectively, what is at stake when one considers whether an agent is programmatically or episodically autonomous. Admittedly, in between these two questions there are many other possible questions of varying degrees of generality that (some) agents could ask themselves – e.g. ‘what line of work do I want to get into?’<sup>48</sup> Furthermore, thinking these two levels as disconnected is not useful insofar as any programmatic insight on how I want to live my life should also allow me to ask and answer more concrete questions – e.g. ‘should I accept this particular job

---

<sup>47</sup> Meyers, *Being yourself*, 8.

<sup>48</sup> *Ibid.*



now?<sup>49</sup> In light of these two points, Meyers explains that there is a *degree* difference between episodic autonomy and (different forms of) programmatic autonomy. Autonomy, therefore, admits degrees, e.g. an agent could exhibit episodic autonomy, or “narrow” programmatic autonomy, without being fully programmatically autonomous.<sup>50</sup>

Throughout my thesis, I accept and defend the idea that autonomy is not an all-or-nothing matter. This claim has at least three implications: first, one can have more or less control over one’s general life – indeed, being in *full* control of one’s life seems at odds with human social life as we know it; second, one may be more or less autonomous depending on which aspect of one’s life is at stake (e.g. one’s sexuality, one’s family life, one’s career); and third, in the case of ‘intersectional’ identities, membership to different groups may allow for different degrees of autonomy in different contexts (e.g. one may occupy a position of privilege because of one’s race or class but not because of one’s gender).<sup>51</sup>

Nonetheless, I also argue that for the particular purpose of limiting one’s oppression, episodic autonomy on its own should not be considered enough. As it will become clearer in the chapters that follow, limiting one’s oppression requires both a significant control over relevant aspects of one’s life and minimal awareness of the “programmatic aims” which orient one’s life.

---

<sup>49</sup> Indeed, focusing on the autonomy of choices or acts is not *necessarily* incompatible with providing a more global account of autonomous agency. For example, Westlund argues that, even if her focus is on choice and action, it is possible to draw from this definition what an autonomous agent or an autonomous life is: an autonomous life would be the result of exercising capacities for choosing or acting in an autonomous way for an extended period of time. (Westlund, “Rethinking Relational Autonomy,” 28)

<sup>50</sup> Meyers, *Being yourself*, 8.

<sup>51</sup> The term ‘intersectional identity’ is used to refer to multiple, complex or *mestiza* identities which are intersected by different race, class, ethnicity, gender, and sexual orientation norms. The term has been used both to analyse the way in which individuals may be “multiply oppressed” and to signal the opportunities for resistance available to intersectional identities. (See for example María Lugones, *Pilgrimages/Peregrinajes: Theorizing Coalition Against Multiple Oppressions* (New York: Rowman & Littlefield Publishers, 2003) For example, some argue that those intersected by different forms of oppression or subjected to (possibly conflicting) social and cultural injunctions may have critical perspectives vis-à-vis (some of) their identities more readily available to them than those “unified selves” who satisfy the requirements of hierarchical accounts (explained below). See for example Diana Meyers, “Intersectional Identity and the Authentic Self? Opposites Attract!” in: Meyers, *Being yourself*; see also Edwina Barvosa-Carter, “Mestiza Autonomy as Relational Autonomy: Ambivalence & the Social Character of Free Will,” *The Journal of Political Philosophy* 15, no. 1 (2007).

Going back to the generalities of procedural accounts, the different alternatives which I discuss below share another key feature: they aim at ‘content-neutrality’.<sup>52</sup> What this means is that, as long as formal corroboration procedures are carried out, the *contents* of what could be autonomously valued, preferred or desired are (purportedly) not limited by anything other than the *agent’s* perspective. The matter for a proceduralist would not typically be *what* I can autonomously desire but to clarify what autonomously desiring *anything* means. For example, desiring something that most people would deem harmful, unethical or uninteresting does not jeopardize one’s autonomy vis-à-vis a desire as long as the desire is “one’s own”. This feature, as I explained in the Introduction, is typically used to distinguish procedural accounts from the ‘substantive’ accounts which I analyse in Chapter 4.

One important qualification is nonetheless needed. The fact that procedural models do not explicitly associate autonomy with specific preferences or lifestyles does not exclude that procedural accounts might (more or less) indirectly make such associations. For example, in procedural terms autonomy is compatible with having preferences for traditional “feminine” roles. Still, proceduralists might end up favouring traditional “masculine” characters if they bring into their models, say, definitions of ‘reflective competence’ which presuppose ideals of rationality which have been long deemed by feminists as “masculine” (I say more about this in Chapter 2). In a word, it is questionable whether content-neutrality automatically secures complete *value*-neutrality.

Indeed, some autonomy theorists point out that there are two ways in which (procedural or substantive) autonomy theorists may incorporate normative content into their accounts: (i) by explicitly restricting the contents of what can be autonomously chosen, valued, desired or done; (ii) through the normative standards presupposed by the conditions

---

<sup>52</sup> See John Christman, “Autonomy and Personal History,” *Canadian Journal of Philosophy* 21, no. 1 (Mar., 1991), 22; Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press, 1988), 29; Alfred Mele, *Autonomous Agents*. (Oxford: Oxford University Press, 2001), 160.

or procedures set by a model of autonomy. As I explained above, the first strategy is explicitly disavowed and generally avoided by procedural theorists. However, insofar as procedural theorists outline validation procedures and set conditions which agents would need to meet to be candidates for autonomy, they cannot avoid *ii*. As Fabian Freyenhagen notes, “even demanding that a decision-making process is such that it does not involve any manifest inconsistencies, is to import normative content: consistency is just one norm among others, and one that is contestable, as, for example, the possibilities of true moral dilemmas and of moral conversions suggest.”<sup>53</sup> Moreover, as I discuss in Chapter 2, a model of autonomy could even presuppose “ethical” values through some of its competence conditions.

To capture the abovementioned issues, Meyers distinguishes two different axes which could be used to classify autonomy models: a ‘Directivity Axis’ which “reveals whether an autonomy theory preempts or honors the judgment of individual agents” and a ‘Constitutivity Axis’ which “lays bare the normative gears that drive competing accounts of autonomous choice and action.”<sup>54</sup> According to Meyers, a theory may be content-neutral on the Directivity Axis but still be “value-utilizing” on the ‘Constitutivity Axis’.<sup>55</sup> Meyers argues that “it is difficult, perhaps impossible, to conceive an account of autonomy that is not value utilizing” insofar as, as it was signalled above, values are required to define decision-making processes and necessary sets of agentic competences.<sup>56</sup> Throughout this thesis I frequently signal how different procedural models (implicitly or explicitly) “utilize” values in spite of aiming at content-neutrality. In Chapters 2 and 3, I signal the dangers of not explicitly acknowledging the values presupposed by a model of autonomy which aims at limiting oppressive interferences.

---

<sup>53</sup> Fabian Freyenhagen, “Autonomy’s Substance,” *Journal of Applied Philosophy*, 34, no. 1 (February 2017), 120. Meyers makes a similar point, she claims: “even rational choice theory—a paradigm of value neutrality—rests on a constitutive set of epistemic values including consistency, transitivity, and knowledge of pertinent facts.” (Meyers, “Feminist Debates Over Values in Autonomy Theory,” 120).

<sup>54</sup> *Ibid.*, 115.

<sup>55</sup> *Ibid.*

<sup>56</sup> *Ibid.*, 120.

Going back to my main discussion, the idea that one's desires or character traits could *not* be "one's own" might appear counterintuitive and requires unpacking. To make sense of this notion it is important to note that the mere *fact* of "having" a certain desire or "showing" a certain trait of character is not enough to determine one's "authority" vis-à-vis that desire or trait. This claim is justified and illustrated in the literature by considering (at least) three kinds of cases:

The first cases are those which involve some sort of "inner conflict". Imagine for example that I have a certain desire which I rather I did not have. My having this desire *in spite of* my "inner policy" to reject or condemn this desire (or this kind of desire) would typically indicate my 'alienation' and, in certain models discussed below, my heteronomy vis-à-vis this desire. (Let's grant for now that the kind of rejection just described amounts to alienation and that it would suffice to make me heteronomous.) In this case, I would be heteronomous even without the intervention of others and just because I fail to direct myself in light of the set of values and commitments which define the kind of person who I *want* to be.

'Hierarchical' accounts of autonomy (discussed below) give these "inner policies" a special (higher) status in the agent's motivational structure. Namely, according to hierarchical accounts, these "inner policies" bear a special and close relation to *who* the agent *is* and, therefore, "inner policies" can be used to assess whether agents are autonomous or heteronomous vis-à-vis other desires or values. When this hierarchical assessment takes place without considering anything more than the *structure* of the agent's psychology at the time when the reflective test of autonomy takes place, the strategy adopted is labelled 'internalist'. Internalist accounts have been criticized both by other proceduralists and by substantive theorists. In this chapter, I present the externalist procedural critique of internalist accounts.<sup>57</sup>

---

<sup>57</sup> I deal with the substantive critique of (internalist and externalist) procedural accounts in Chapter 4.

Externalist procedural theorists argue that one's dispositions could *not* be one's own as a result of *external* interferences or interventions and, crucially, as a result of the interventions of *other* agents.<sup>58</sup> These interferences or interventions could be more or less easy to identify. On the more "evident" side of the spectrum, we could locate those cases which involve outright manipulation. For example, an agent under hypnosis (which she has not previously consented to) is not the "source" of her actions or dispositions in the same way that a non-manipulated agent doing exactly the same things is. Indeed, we might be tempted to say that this agent has been reduced to a puppet-like performance and that whoever "pulls the strings" is the one governing the action. Moreover, even if *while hypnotised* the agent approved of her situation, her actions or her dispositions, we might understandably be suspicious of a consent or endorsement given in these conditions.

As I explain below, (externalist) procedural theorists typically resort to 'procedural independence' conditions to deem cases involving manipulation or deception 'heteronomous'. Procedural independence conditions restrict the subjective and intersubjective conditions under which a reflective test of autonomy could be valid and intend to secure sufficient "independence of mind" to carry out an autonomy test. Agents can lack independence of mind for different (external) reasons, for instance, as the hypnosis example above illustrates, one can lack independence of mind because one cannot judge in a non-manipulated manner. However, one can also lack independence of mind without being

---

<sup>58</sup> See for example Alfred Mele's "psychological twins" counterexample against internalist accounts. (Mele, *Autonomous Agents*, 145). Mele also grants the existence of non-intentional external interventions which may block one's autonomy – e.g. one may change one's value framework as a result of a brain disorder. In cases like this, Mele argues, non-autonomy will be different from heteronomy. (Ibid., 168-9). The effects of "accidental" interferences are not the focus of this thesis. The reason for this is that oppression typically involves "intentional" interventions (in the sense of "purposive" interventions) – e.g. as I explained in the Introduction, oppressive interventions typically aim at fixing certain groups in positions of subordination or privilege. In Chapters 2 and 3 I argue, however, that this "purposiveness" does not necessarily involve the actions of individual manipulators or deceivers, and that "interferences" can be structural or institutional. Moreover, others might reinforce or contribute to one's oppression in spite of their "good intentions". I still use the term 'heteronomy' to refer to cases where no individual "other" can be made responsible for one's oppression.

coerced *at present*, which takes me to the third sense in which one's values or desires could fail to be one's own:

Agents can act, desire, and have preferences as a result of *past* interventions or interferences which might call into question the extent to which present actions or dispositions are truly the agents'. Procedural theorists sometimes turn to (rather farfetched) examples which involve some form of "programming" through brainwashing or neurosurgery to make the point that one's motivational structure might not be one's own because of its *origin*.<sup>59</sup> Leaving aside the artificial examples that are sometimes used in the literature to illustrate this point, what this worry captures is that an agent's motivational structure might be "freely held" at present and be, nonetheless, "coerced". Considering *how* a motivational structure was formed might be relevant to determine an agent's autonomy vis-à-vis her values and motivations. As I discuss below, *historical* procedural models incorporate different conditions into their models to address this worry.

Crucially, the worry that something about the genesis of a value might jeopardise an agent's autonomy merits serious consideration in the context of a discussion on the possibility of autonomy under oppression. While manipulation by an evil neuroscientist may appear an artificial way to think about common forms of oppression, it is worth considering if the model of manipulation (understood broadly) could shed light on some of the ways in which an oppressive context may be autonomy-undermining. For example, it might be possible to argue that certain forms of value indoctrination which happen under some forms of oppression may be as bad as the abovementioned "programming". The reason for this, a procedural argument could go, is that the value framework acquired as a result of indoctrination might be impossible to assess without an "external" perspective (i.e. one that does not assume the values under assessment) and might be extremely difficult to change or

---

<sup>59</sup> See *Ibid.*, Chapter 9, for cases involving "psychological cloning", magically creating agents, being affected by the magnetic forces of the Bermuda Triangle, amongst others.

modify. The issue with this argumentative strategy, however, is that the same might hold true for many of the values or traits of character that we acquire through socialisation in general (i.e. oppressive or not). Indeed, some may even argue that “standard child-rearing techniques” sometimes get quite close to those techniques feared by proceduralists.<sup>60</sup>

Could we distinguish between different autonomy-enhancing and autonomy-undermining forms of socialisation in procedural terms? We sure could identify some scenarios which appear incompatible with autonomy as proceduralists picture it. For example, an upbringing in which one is denied minimal education (broadly understood) could be thought as “objectively” autonomy-undermining in procedural terms, since one would lack the minimal reflective competence needed to carry out the necessary reflective self-assessment to determine one’s autonomy procedurally. However, it is difficult to go further than this when limiting “illegitimate” forms of socialisation without compromising procedural ‘content-neutrality’. Let me unpack this claim:

As I will show below, historical procedural theorists find themselves in a tricky position. Namely, they need to find a balance between setting external conditions for the self-assessment associated with autonomy *and* guaranteeing that there is still a place for the perspective of the “self” who attempts to govern herself. The question is how far we should go when setting limits to the kinds of histories and lives compatible with autonomy. This is probably a question that haunts all theorists of autonomy. Typically, the fear is that, when answering this question, one might end up unjustifiably, arbitrarily or paternalistically valuing certain modes of life over others. Moreover, in the context of a discussion of oppression, we should consider a further danger, namely that of making oppression worse by disregarding

---

<sup>60</sup> See for example: “It is perhaps worth recalling that ‘standard child-rearing techniques’ include such processes as operant, aversive, and classical conditioning; role model imitation; blind obedience to and subsequent internalization of behavioral norms; uncritical acceptance of propositions on the authority of parents and teachers; and so on.” (Robert Noggle, “Autonomy and the Paradox of Self-Creation: Infinite Regress, Finite Selves, and the Question of Authenticity,” in: James Stacey Taylor, *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy* (Cambridge: Cambridge University Press, 2005), 102-3. I say more about Noggle’s view later in this chapter.

the experiences of the oppressed, their resistance strategies, and their own perspectives on their own lives.

With the above in mind, I consider crucial to distinguish ‘*orthonomy*’ (i.e. government by “right” norms or reasons) from *self-government*.<sup>61</sup> I will therefore assess different models of autonomy in light of their capacity to make this distinction. Making this distinction does not preclude identifying social situations which are better or worse – indeed, I do identify oppressive circumstances throughout my thesis and this implies the judgement that something is *wrong* about these cases. However, looking into forms of oppression to decide what it would mean to be less heteronomous under these conditions (e.g. how agents could gain opportunities to limit, in their own terms, heteronomy in these conditions) is different from making autonomy incompatible with oppression. In a word, my point is that it should generally remain possible to be autonomous in circumstances that others would find problematic and even to autonomously hold values that others would have good reasons to reject – this even applies to values that others might deem “oppressive”.<sup>62</sup>

In light of the above, I argue that John Christman’s historical solution is particularly promising. What distinguishes Christman’s account from other historical accounts is that for Christman, as long as the agent is minimally competent, different life developments become materials that inform a reflective process that is carried out *from the agent’s perspective*. What is at stake in Christman’s model is therefore not if a certain character trait has an “objectively

---

<sup>61</sup> See Benson for a distinction between ‘self-government’ and ‘orthonomy’. While for Benson the former is “the power to *take ownership* of one’s actions”, the latter is defined as “the power to *get things right* or the ability to adopt the preferences or values one ought to have (or at least to avoid those one ought not to have)” (Paul Benson, “Feminist Intuitions and the Normative Substance of Autonomy,” in Taylor, *Personal Autonomy*, 132.) See Tom O’Shea, “Autonomy and Orthonomy,” *Journal of Moral Philosophy* 12, no. 5 (2015) for an argument against the view that orthonomy (as government by “objective” standards of reason) necessarily implies “a missing self” – i.e. an “unmooring” from the self who is supposed to do the governing in self-government. (Ibid., 627). O’Shea argues that most models (even those who explicitly object orthonomy) adopt some ideal of “good” rationality. As it will become clearer in Chapter 2, I agree with the latter point. However, I maintain that the distinction between ‘self-government’ and ‘orthonomy’ should be kept insofar as self-government must remain compatible with living life by different values and in a wide variety of normative contexts (including non-ideal ones). I develop this point in Chapter 4.

<sup>62</sup> I say “generally” because, as it will become clear in Chapters 4 and 5, I do propose that a minimum of ‘non-domination’ must be met for autonomy to be decided in procedural terms.



good” history (or conversely – as Mele puts it – if the said trait lacks an authenticity-incompatible history)<sup>63</sup> but whether the agent could (putting it roughly, to avoid technicalities for now) “approve of” that trait given *her* value framework *even* if she became aware of the trait’s history. By framing autonomy in this way, Christman is particularly successful in avoiding restricting what can be autonomously chosen by competent and non-manipulated agents. Additionally, as I argue below, Christman’s model is promising in that it gives agents the opportunity to question the effects of oppression on themselves, insofar as it asks that agents could “approve of” aspects of themselves *even* if made aware of the historical and social elements which have shaped their personalities.

To summarize, in this chapter I argue that Christman’s account is promising for achieving three desiderata: (a) avoiding paternalism and perfectionism as much as possible; (b) stopping short of ascribing autonomy in implausible cases; and (c) starting a discussion on the necessary reflective materials that could make the perspectives of oppressed agents critical enough to limit their oppression. The proceduralism of Christman’s theory is particularly suitable for achieving (a), while its historical dimension helps with (b) and (c) and makes his model superior to purely internalist procedural approaches and to other historical accounts (of which I provide examples in the next section)

The question that remains unanswered in this chapter is, however, whether the model of “manipulation” or “deception” which frames most of the attempts to define procedural independence (including Christman’s), is useful, all things considered, to think about oppression. I focus for now on identifying those features in historical procedural models that are promising to deal with oppression. I then problematize, in the chapters that follow, the underlying notion of oppression presupposed by available historical procedural models.

---

<sup>63</sup> Mele, *Autonomous Agents*, 172.

## II. 'Internalist' Proceduralist Models and Their Limits

Before discussing historical models, it is helpful to get a purely 'internalist' model in view. The account often ascribed to Harry Frankfurt is a good example of an internalist account. Even if Frankfurt's interest was to analyse freedom of the will in the context of debates about determinism (rather than autonomy in the context of liberal rights), his definition of the person has been traditionally incorporated into theories of autonomy.<sup>64</sup> In a Frankfurtian theory, autonomy is decided by creating a hierarchical distinction between lower-level and higher-level volitions and asking for endorsement, through a second order volition, of a first order volition. Frankfurt claims:

Someone has a desire of the second order either when he wants simply to have a certain desire or when he wants a certain desire to be his will. In situations of the latter kind, I shall call his second-order desire "second-order volitions" or "volitions of the second order". Now it is having second-order volitions, and not having second-order desires generally, that I regard as essential to being a person.<sup>65</sup>

So, persons (or agents) do not simply "feel" instincts or desires: it seems characteristic of a person to be able to problematize them, taking them as objects of reflection. Higher-order volitions "sanction" or endorse lower-level ones and through this sanctioning the agent commits wholeheartedly to them and makes them her will. In the literature on autonomy, this possibility of reflectively sanctioning, validating or, more generally, "coming to terms" with lower-level contents is referred to as the 'authenticity' condition.

---

<sup>64</sup> See Frankfurt, "Freedom of the Will and the Concept of a Person."

<sup>65</sup> Harry Frankfurt, *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988), 16.

However, this account of autonomy could have counterintuitive consequences. One common way of criticizing hierarchical accounts is by signalling the risks of a regress – i.e. we can always ask if higher-order values are themselves authentic and on what grounds they are so. If we rely on the same ‘hierarchical’ strategy to validate higher-order values (i.e. we resort to “even-higher-order” values), then the same question arises again. I do not focus on this objection here partly because it has been sufficiently discussed, and partly because I believe that this problem is not so pressing for the lives and values of *actual* agents.<sup>66</sup> That is: I am willing to accept (with others) that autonomy implies (among other things) the adoption of “inauthentic” values – i.e. the adoption of values which are not “originally” the agent’s and which cannot be validated from an external or higher-order perspective.

For example, Robert Noggle claims that, at very early stages of our finite social lives, we acquire “core” values “inauthentically”, that is, through the interventions of others during socialisation and upbringing.<sup>67</sup> Therefore, deeming agents heteronomous based on the “inauthenticity” of this value-acquisition would mean making autonomy incompatible with any form of socialisation. According to Noggle, core values do progressively *become* “authentic” but this does not happen thanks to an external perspective which could allow agents to independently assess core values or the processes that gave place to those core values.<sup>68</sup>

Leaving aside potential issues with the notion of authenticity itself (which I discuss in Chapter 3), I find strategies like Noggle’s fruitful if we are to understand autonomy in a way which is compatible with a relational and social conception of agency. To put it shortly for now, my point is that autonomy needs to be thought as a task achievable through materials, values, and character traits which are (in many important respects) “given” to us.

---

<sup>66</sup> See Taylor’s complete account of the regression objection raised against hierarchical models in Taylor, *Personal Autonomy*, 6-10.

<sup>67</sup> Noggle, “Autonomy and the Paradox of Self-Creation,” 95.

<sup>68</sup> *Ibid.*, 99-100.

Furthermore, what interests me particularly in the context of a discussion of autonomy under oppression, is to consider if we could accept the “tainted” origins of some of the materials with which we make ourselves agents and proceed from there with a view to limiting heteronomy. The externalist objections and strategies which I consider in what follows are compatible with this strategy.

Going back to my discussion of internalist accounts, these strategies are also criticized for failing to make any reference to the context or circumstances in which higher-order endorsement takes place. Externalists note, for example, that an agent proceeding while under (non-consented) hypnosis would be as autonomous as one who reflects freely.<sup>69</sup> To give different treatment to cases like these, externalists claim, one must ask for more than authenticity and add conditions which define *legitimate* circumstances for reflection and endorsement.

In addition to authenticity, proponents of externalist-procedural models also typically connect autonomy to a second requirement: that the agent has a capacity to reflect competently and independently. Proponents of externalist-procedural models note that, to be able to say that certain contents are “one’s own”, not only internal critical distance vis-à-vis these contents is needed (e.g. to take a content as an object of reflection) but also a minimal degree of external independence of mind.

In the hypnosis scenario mentioned above, for example, it is difficult to see how the outcome of a reflective process could be considered meaningful to the agent’s autonomy. Instead, one would intuitively claim that the agent is not in a position to make an autonomous decision, or that there are reasons to be sceptical about the outcome of the reflective process until one sees if it coincides with a test carried out under circumstances where the agent is in charge of her reflective capacities.

---

<sup>69</sup> I leave the door open for it to be compatible with autonomy that there can be cases in which hypnosis was arranged by the agent herself (to overcome a phobia or an uncontrollable desire, for example).

Gerald Dworkin, like Frankfurt, associates autonomy with a capacity for higher-order assessment of lower-level preferences or desires. For Dworkin, critical distance from one's dispositions is crucial for autonomous agents because “[b]y exercising such a capacity we define our nature, give meaning and coherence to our lives, and take responsibility for the kind of person we are.”<sup>70</sup> However, this is not the whole picture of autonomy:

To overcome the difficulties associated with internalist accounts à la Frankfurt, Dworkin incorporates into his own account a condition of ‘procedural independence’. Procedural independence plays a more fundamental role than other conditions set by externalist-procedural models (e.g. authenticity) insofar as it is a *condition of possibility* for any meaningful and trustworthy reflective validation of preferences or desires. That is: the eventual correspondence between a self (a certain “kind of person”) and the contents under analysis needs to be established or verified under *appropriate* reflective conditions.

Procedural independence is meant to help us distinguish between legitimate and illegitimate ways of affecting agents’ capacities and psychological structures – namely between merely “influencing” people’s higher order judgements and “interfering” with them. Procedural independence should therefore not be thought as asking for radical independence of mind (if such a thing is even possible). Dworkin indeed acknowledges that shaping agents’ reflective capacities through education or socialisation may indeed “promote and improve” these capacities.<sup>71</sup> Dworkin claims:

With respect to autonomy, conceived of as authenticity under conditions of procedural independence, the paradigms of

---

<sup>70</sup> Dworkin, *The Theory and Practice of Autonomy*, 108. Dworkin’s hierarchical theory had two fundamental moments: one in which it was crucial for a subject to be able to *identify* with her second-order contents, and another one (the position reported here) in which the key is to be able to *raise the question* of whether one wants to identify with certain contents or wants to attempt to change them. He argues: “It is not the identification or lack of identification that is crucial to being autonomous, but the capacity to raise the question of whether I will identify with or reject the reasons for which I now act.” (Ibid., 15, emphasis added).

<sup>71</sup> Dworkin, *The Theory and Practice of Autonomy*, 18.

interference are manipulation and deception, and the analytic task is to distinguish these ways of influencing people's higher order judgments from those (education, requirements of logical thinking, provision of role-models) which do not negate procedural independence.<sup>72</sup>

External procedural models define the *kinds* of interpersonal relations that are compatible and incompatible with autonomy. It is possible to be in an autonomy-undermining relation with others (e.g. the one – or ones – who coerces me or manipulates me). Conversely, it is possible to be in autonomy-compatible relations to others, as long as procedural independence conditions protect the agent from becoming an instrument for *somebody else's* interests or will.<sup>73</sup>

It is necessary to add that what is being safeguarded through the procedural independence condition is not the possibility of making *ideal* or *perfect* decisions, but the *intersubjective* conditions in which a decision can legitimately be considered *one's own*. As it was explained above, what is relevant for proceduralists is *not* to limit the *contents* of what can be autonomously decided or chosen by an agent but to define a procedure that could prima facie be met by any preference, value, or personal trait.<sup>74</sup> An agent who has not been deceived and reasons in a non-manipulated way, can be considered “responsible for his reasoning and his conclusions”.<sup>75</sup> Dworkin's emphasis on responsibility seems to suggest that there is also value for autonomy when an agent rightfully makes *her own* “wrong” decisions. That is: the agent must be in fact assessing motives in a way that is coherent with *her* values. So, procedural independence seems to be protecting individuals not from mistake (as an

---

<sup>72</sup> Gerald Dworkin, “Autonomy and Behavior Control,” *The Hastings Center Report* 6, no. 1 (February 1976), 26.

<sup>73</sup> Gerald Dworkin, “The Concept of Autonomy,” in Christman, *The Inner Citadel*, 59-60.

<sup>74</sup> Dworkin, *The Theory and Practice of Autonomy*, 29.

<sup>75</sup> *Ibid.*, 161.

objective criterion) but from certain others (those who misguide the agent, deceive her, undermine or distort her capacities for reflection, etc.).

But *when* exactly is ‘procedural independence’ achieved? Do we have a positive definition of procedural independence? Dworkin proceeds *negatively*: his strategy is to outline cases in which procedural independence is *not* met – i.e. cases in which: “[a] person’s motivational structure may be *his*, without being his *own*”.<sup>76</sup> The first one of these cases was already mentioned above and concerns the possibility of having been (or being at present) manipulated or deceived: an agent is *not* in a position of procedural independence when “the identification with his motivations, or the choice of the type of person he wants to be, may have been produced by manipulation, deception, the withholding of relevant information, and so on.”<sup>77</sup> Second, a failure in procedural independence may occur when “[a]n individual may identify or approve of his motivational structure because of an inability to view in a critical and rational manner his situation.”<sup>78</sup>

Dworkin’s definition of procedural independence, however, appears incomplete if one does not define, for example, what exactly counts as “viewing one’s situation in a *critical* manner”. Moreover, it is not exactly clear what qualifies as an illegitimate imposition of “the kind of person” one wants to be (i.e. of higher order values). For example, would the transmission of pernicious stereotypes (e.g. those that seriously erode the confidence necessary for critical self-examination) through “normal” means of socialisation (i.e. through processes that do not involve obvious deception or manipulation) be compatible with procedural independence?

Dworkin does provide some helpful clarifications and classifies different *methods* of influencing people’s motivations and value-frameworks according to how preferable they are

---

<sup>76</sup> Dworkin, “Autonomy and Behavior Control,” 25.

<sup>77</sup> *Ibid.*

<sup>78</sup> *Ibid.*

having autonomy in mind. For example, methods that protect self-respect and dignity, that are non-intrusive and safeguard physical integrity, that are reversible or do not last indefinitely (to allow for changes of mind), or that work on cognitive and affective dimensions of the subject instead of “short-circuiting” desires and beliefs, tend to preserve an agent’s capacity for autonomy. On the other hand, methods that are destructive of one’s ability to think rationally, that affect the personal identity of agents, or that rely on deception should be avoided.<sup>79</sup>

However, what one cannot find in Dworkin’s theory is an account of *how* exactly a procedural model of autonomy should incorporate the abovementioned clarifications. His theory is useful negatively (i.e. to show the limits of internalist accounts) but does not offer a developed alternative.

In light of the above, historical procedural models emerge as strong candidates to both (i) successfully deal with the insufficiencies of hierarchical accounts; and (ii) do so by *explicitly* defining a reflective procedure that is to be (or could be) carried out by the same agents whose autonomy is the question.

The models analysed in the next section could be considered to be developing more explicitly and to a fuller extent external (and even historical) elements that are present but somewhat underdeveloped in Dworkin’s theory. That is: Dworkin’s model could be considered a historical one (albeit not explicitly) insofar as it is not indifferent to what happens *before* the actual process of reflective self-assessment. However, I will especially focus on externalist models that (in addition to including a “history-sensitive” procedural independence condition à la Dworkin) are historical in a stronger sense. Namely, some historical models require that the *agent* somehow reflectively considers the history of the contents under analysis.

---

<sup>79</sup> Dworkin, “Autonomy and Behavior Control,” 27-8.



In the next section, I analyse the most influential procedural historical accounts of autonomy and (as I have already anticipated) I defend the position that Christman's model should be favoured over other historical models insofar as it provides valuable tools to limit oppressive interventions *from the agent's perspective*.

### III. Historical Models

Since his early reflections on autonomy, Christman has been mainly concerned with accounting for self-government *among* social and political influences. Indeed, Christman rejects ideals of radical self-authorship – he claims: “no person is self-made in the sense of being a fully formed and intact ‘will’ blossoming out of nowhere”.<sup>80</sup> In light of this fact, it becomes central in Christman's theory to define in which (or after which) *external* conditions autonomy is possible. Moreover, Christman also aims at elucidating, through a theory of autonomy, *when* agents' values or preferences are morally and politically meaningful:

What is needed, [...], is to establish an account of self-determination or autonomy that would help determine just when and if the values and preferences we find ourselves with deserve the centrality that moral and political theories place on them.<sup>81</sup>

If a model of autonomy considers *every* preference or value as being *necessarily* an expression of the agent's authority over herself, then it fails its fundamental task which should be exactly figuring out *which* preferences or values should be valued *qua* expressions of the agent's self-determination. Christman's interest is ultimately to put in place a procedure that allows us to decide when self-government is not unduly affected by social influences such that liberal

---

<sup>80</sup> Christman, “Autonomy and Personal History,” 1.

<sup>81</sup> *Ibid.*

rights should be granted (or, in the opposite case: when paternalistic interventions could be authorised). So how is this procedure defined?

Christman's historical model has changed quite significantly since its first elaboration in 1991. In early versions of the model, Christman's focus is on analysing whether *desires* are autonomous and the test for determining the autonomy of a desire consists in whether the agent would have resisted the *origin or history* of such a desire had she been aware of it.<sup>82</sup> (The counterfactual nature of the condition is meant to avoid an overly intellectualist, rationalist account of autonomy – i.e. Christman is willing to concede autonomy even to agents who have not actually carried out a reflective test of autonomy but could pass it if they were to carry it out.) Christman's key move is to give a vital explanatory role to the *origin* of the personal features under scrutiny: one can only assess personal features by “essential reference” to influences which shape one's personal development “throughout our personal histories”.<sup>83</sup> According to Christman, structural ‘time-slice’<sup>84</sup> approaches (like Frankfurt's) are ill-equipped to account for autonomy because these models overlook the essential, that is: “the manner in which the desire was formed - the conditions and factors that were relevant during the (perhaps lengthy) process of coming to have the value or desire”.<sup>85</sup>

In Christman's most developed and systematized account of autonomy, *The Politics of Persons*, the aspects of “selves” under assessment should be understood broadly enough to include not merely specific desires or preferences but also more general and stable value-

---

<sup>82</sup> The original model could be summarized as follows: A person is autonomous regarding a *desire* if: a) the person did not (or would not have) resisted the process that led to that desire; *and* b) this (actual or counterfactual) lack of resistance is not explainable by the fact that some factors inhibit self-reflection; *and* c) reflection is minimally coherent and without self-deception. So, Christman's focus at that stage of his model was on the *process* of development of the desire and not on the desire itself. (Ibid., 11).

<sup>83</sup> Ibid., 1.

<sup>84</sup> That is: models that consider the psychological structure of the agent at a particular point in time. Christman's model differentiates from time-slice approaches because it considers the agent *diachronically*. However, it might still be said that the reflective process (in Christman's most developed model at least) is performed in a time-slice, because the agent would reflectively validate her character or organizing values at a particular point in time, even if she considers her biographical/historical information. In fact, it might be even possible to assess oneself differently in different moments of one's life.

<sup>85</sup> Ibid., 10.

orientations and character traits. Christman explains that “room should be made for asking about our autonomy relative to the most basic value commitments that ground our identity”.<sup>86</sup> From now on, following Christman’s use of the term, I will refer to these more general aspects of oneself as ‘organizing values’.

Moreover, currently Christman’s model does not take as an object of reflection historical processes *per se*. That is, the historical process leading to a content is not something agents should accept or not, but (actual or counterfactual) awareness of this historical process provides (or could provide) an opportunity to gain a special understanding that leaves agents better positioned to judge the *content* itself. Christman states: “The object of the reflection here, it should be stressed, is not the process of character formation itself, but the trait such formation produces *in light of that development*”.<sup>87</sup>

This constitutes a difference in emphasis compared to other historical models of autonomy in which the assessment of a trait of character or disposition considers the historical process of character formation more directly. This is the case in Alfred Mele’s account. I compare Mele’s and Christman’s accounts briefly to distinguish between two different forms of incorporating historical conditions into procedural accounts.

Like other externalists, Mele also argues that an agent’s psychological structure, on its own, does not provide sufficient information to assess an agent’s psychological autonomy. This claim is supported by turning to “psychological-twins” thought experiments: Mele describes cases which involve two agents with identical psychological structures (i.e. they have the same preferences and values and the same attitudes vis-à-vis these preferences and values) and asks us to imagine that only one of them was (against her will) programmed or

---

<sup>86</sup> Christman, *The politics of persons*, 148. From now on, my general references to “Christman’s model” or “Christman’s account” should be understood as referring to Christman’s developed and systematized view in *The Politics of Persons*. I acknowledge that Christman has revised some particular points of his model, these revisions are considered in Chapter 3, as potential replies to the worries I raise vis-à-vis Christman’s account.

<sup>87</sup> *Ibid.*, 146; emphasis in the original.

brainwashed to have the psychological structure she now has. Mele argues that we would not intuitively deem the brainwashed agent autonomous and that this intuition cannot be explained from a purely internalist perspective. Mele claims:

... given the psychological similarities between the two agents, the difference in their current status regarding autonomy would seem to lie in how they came to have certain psychological features that they have, hence in something external to their here-and-now psychological constitutions. That is, the crucial difference is historical; autonomy is in some way history-bound.<sup>88</sup>

In light of the above, Mele adds external conditions which need to be met for ‘psychological constitutions’ to count as ‘authentic’, which is a precondition of autonomy. Authenticity in Mele’s account is understood as a “*historical* property of agents required for responsibility for the possession of a pro-attitude” (values or preferences are examples of ‘pro-attitudes’).<sup>89</sup> An agent is responsible for the possession of a particular pro-attitude only during the intervals where one can say that she is *not compelled* (unless the compulsion was arranged by the agent herself) to *possess* this pro-attitude.<sup>90</sup> One is compelled to possess a pro-attitude when the attitude was acquired in a way that bypassed the agent’s control over her mental life (e.g. the ability to manage one’s desires according to judgements) *and* this bypassing results in the agent’s *inability to shed* the pro-attitude.<sup>91</sup>

For Mele, then, authenticity (and psychological autonomy) are neither automatically lost when one’s values or preferences are influenced by others nor when one is “stuck” with certain pro-attitudes. What decides inauthenticity is whether any of the circumstances just

---

<sup>88</sup> Mele, *Autonomous Agents*, 145-6.

<sup>89</sup> *Ibid.*, 166; emphasis added.

<sup>90</sup> *Ibid.*

<sup>91</sup> *Ibid.*, 167. ‘Shedding’ a pro-attitude includes an ability to “eradicate” or to “significantly attenuate” the said pro-attitude. (*Ibid.*, 153).

described is the result of an *authenticity-incompatible* history –e.g. brainwashing or manipulation.<sup>92</sup> In addition to the condition that one cannot be compelled to possess one’s motivational structure, Mele adds two additional conditions for autonomy: an agent should hold beliefs that lead to informed deliberation, and she should be a “reliable” deliberator.<sup>93</sup> I say more about Mele’s competence conditions in Chapter 2.

According to Mele, his definition is successful in accounting for illegitimate interventions *without* making autonomy incompatible with “normal” processes of socialisation. Why? Because socialisation typically happens when there are still no reflective capacities to bypass (e.g. learning to assess and control one’s desires or preferences is also a part of socialisation). Moreover, when socialisation is non-dogmatic (e.g. when it is different from indoctrination) an agent is also typically given tools that could eventually allow her to modify or change the preferences or values acquired during socialisation.<sup>94</sup>

To be sure, not all the preferences or values that we acquire through usual forms of socialisation will be easy to change (indeed, this seems to be at odds with the way in which actual agents relate to their values). Nonetheless, Mele’s point is that, from an externalist point of view, it is very different to be attached to values acquired through non-coercive means than to be “tricked” into these values. Only in the latter case does “attachment” to one’s preferences or values become problematic according to Mele. Ultimately, the key difference in Mele’s account is a historical-evaluative one: some disposition-formation *processes* are incompatible with autonomy.

In Mele’s view, then, there are two kinds of ‘external’ requirements for psychological autonomy: agents need to be competent *and* the history of their pro-attitudes cannot be authenticity-incompatible (as defined above). These requirements are external because they

---

<sup>92</sup> Ibid., 158-9.

<sup>93</sup> Ibid., 187.

<sup>94</sup> Ibid., 167

are not decided by the agents whose autonomy is under assessment themselves (or by others with *mere* reference to the agent's perspective). For example, when a value has a "negative" history, then an agent cannot be autonomous vis-à-vis this value, no matter her reflective or interpretative attempts to assess it. Indeed, Mele claims that the viewpoints or preferences of agents subjected to previous manipulation or "programming", even those of agents who have been made aware of their past manipulation or "programming", should not be taken into account to decide their autonomy: "[w]hat blocks the inference from her reflective preference to autonomy", Mele Claims, "is the etiology of that preference, an etiology that renders a certain pivotal collection of pro-attitudes practically unsheddable."<sup>95</sup>

Christman also includes external conditions in his model. In Christman's case, however, such conditions are restricted to determining whether an agent possesses 'reflective competence.' The latter is defined as an ability to reflect without constriction, pathology or manipulation and thanks to certain enabling conditions (e.g. education).<sup>96</sup> Competence conditions certainly make some personal histories typically incompatible with autonomy. For example, cases of extreme deprivation, or of lack of minimal education or socialization, would make the acquisition of minimal reflective competence extremely difficult. However, once this minimum is guaranteed, agents could potentially endorse *any* content, no matter its history. Nothing impedes, for example, to be autonomous with regards to contents that one was neither free to choose nor is free to *change*, provided one meets the other requirements set by Christman's model.<sup>97</sup> This possibility emerges quite clearly in the following paragraph:

---

<sup>95</sup> Ibid., 170-1.

<sup>96</sup> Christman, *The politics of Persons*, 147. Note that Christman refers basically to the same conditions of 'procedural independence' defined above.

<sup>97</sup> Ibid., 160. Christman claims: "a general ability to do otherwise, or even to desire otherwise, is not in itself a plausible requirement for autonomy. [...], we are often in states that are not revisable by us without tremendous pain, effort, or even outside assistance. We might be madly in love, for example, or unquestionably devoted to our children, and having such desires, values, and commitments could not be shed without great psychological cost, to an extent easily comparable to the effort it takes to resist the pangs of a heroin addiction."

... historical conditions for autonomy should also not instantiate narrow ideals of the good life, such as requiring that all of our desires actually result from reasoned reflection to be authentic (since most did not for most of us). Rather, *the agent's own perspective on adequate historical processes* by which she came to be how she is should remain paramount in the determination of autonomy for that person, ...<sup>98</sup>

Awareness of the historical processes at the origin of one's organising values is meant to leave *agents* in a better position to assess *themselves* these traits – i.e. whether a value is acceptable given its history is for Christman a matter of *subjective* assessment. Indeed, as long as one's history is compatible with the development of reflective competence, historical materials will not provide, on their own, any conclusive information on one's autonomy. So, Christman's includes a *non-perfectionist* historical element in his model of autonomy.

All in all, we can see that the two historical accounts considered so far offer quite different historical conditions. While Mele makes autonomy automatically incompatible with certain personal histories (i.e. those which lead to being compelled to possess certain dispositions or values), Christman does not automatically decide an agent's autonomy or heteronomy vis-à-vis a disposition based on her personal history exclusively. Instead, Christman asks us to consider whether a test carried out *from the agent's perspective* could succeed in validating a specific organising value *even* considering the value's history. In a word, Christman (unlike Mele) leaves the door open for agents to be autonomous vis-à-vis values with histories that others would see problematic or dubious provided the agent is (or could be) sufficiently aware of the history of these values. Moreover, since in Christman's account what is under assessment is not the history of a value but the value itself, Christman

---

<sup>98</sup> Ibid., 13-4; emphasis added.

allows for agents to be autonomous vis-à-vis values *in spite of* finding the histories of these values problematic (e.g. I could find reasons to come to terms with a value that I know I acquired as a result of an upbringing that I reject).

Admittedly, someone could argue that it makes sense to adopt a perspective like Mele's, insofar as we should protect agents from illegitimate interventions like the ones Mele is concerned about. This worry is reasonable and Christman's model does indeed offer these safeguards through his competence conditions: agents should not be deceived or manipulated *at the time* when self-assessment takes place. However, what Christman does not do is to rule out what kinds of *past* upbringings, personal histories or value-formation processes are incompatible with autonomy. In the context of a discussion of autonomy under oppression, I argue, Christman's perspective is therefore preferable insofar as it appears more flexible to accommodate the possibility of autonomous agency emerging from a wide variety of 'non-ideal' historical scenarios.

Moreover, I believe that a solution like Mele's runs the risk of excessively limiting or directing the kinds of lifestyles or characters compatible with autonomy in a way that may even make Mele slip into an account of 'orthonomy' (defined earlier). Ultimately, personal histories are evaluated mostly in light of whether or not they produce rational agents and competent deliberators of certain kinds. Indeed, according to Mele, past interventions which enhance "deliberative excellence", even those interventions which may "bypass the agent's capacities for control over his mental life", are compatible with autonomy.<sup>99</sup> However, if we want to make room for a plurality of lifestyles and ways of thinking (e.g. if we want to grant that an agent may not value a specific form of "deliberative excellence" above other aims), then we also need to account for the fact that personal histories might not be that easy to evaluate from an objective point of view.

---

<sup>99</sup> Mele, *Autonomous Agents*, 183-4. I say more about Mele's competence conditions in Chapter 2.



Admittedly, a potential objection to my point above (and to Christman's model) should be considered: if history is not assessed *objectively* but is a matter of interpretation from the agent's perspective, the chances of self-deception and bias might be greater than in Mele's model. Two replies to this objection are possible:

First, let me note that even if Christman's model was indeed more susceptible to this danger, this solution is still preferable to deciding too strictly what kinds of personal histories or past circumstances are compatible with autonomy. Orthonomy is, I believe, the greater danger that needs to be avoided in this case.

Second, it is worth mentioning that Christman defends his proposal from this kind of objection by adding the proviso that the history which informs reflection needs to be "minimally adequate".<sup>100</sup> In a word: agents need not be fully aware of their psychological histories in the way self-transparent agents would be, but they should not be deceived with regards to the most important origins of their psychological structures. Christman expands on this criterion in the following way:

The person's conception of how she came to adopt a desire or value commitment need only be minimally adequate as an account of her self-development, in the sense that it be consistent with accepted evidence and known causal sequences. It need not be the final or most basic description of her psychology over time.<sup>101</sup>

One is deceived, then, if one's account of the history of a certain value does not match the reality of what actually happened (as known so far).<sup>102</sup> So, to assess a value or desire

---

<sup>100</sup> Christman, *The Politics of Persons*, 154.

<sup>101</sup> Ibid.

<sup>102</sup> Because, as Christman acknowledges, in order to be autonomous, one is not expected to know the "most final" description of one's history. I interpret this to mean that knowledge of new events could eventually come up and one's interpretation of one's values in light of this new unveiled history might change in consequence. For Christman, this would not mean, I believe, that one was deceived before as long as one's partial account of one's history was not the result of deception or manipulation – i.e. the previous account matched those facts intersubjectively accepted as adequate, but they were later proven to be insufficient. (Ibid.)

autonomously, we need an adequate understanding of how the value or desire was formed, and this information can be acquired through our *psychological histories* or considering the process of our *self-development*.<sup>103</sup>

To be sure, the proviso described above protects Christman's theory from some potentially problematic drawbacks – e.g. agents would not count as autonomous if they validate their organising values in light of a personal history which is utterly false. However, as I discuss in Chapters 2 and 3, it is necessary to address more explicitly the fact that histories are *interpretative* objects and that not all readings of one's history could equally serve autonomy as I define it.

And how does exactly Christman's test of autonomy work? What needs to be assessed is whether a trait of oneself (which has a certain history) is "acceptable" in light of who we are. In other words, Christman provides his own version of procedural 'authenticity'. I say much more about Christman's understanding of authenticity in Chapter 3. For now, let me just briefly note the key ways in which authenticity is verified (or ruled out) in Christman's model:

First, authenticity implies a certain correspondence between an organising value and "our narrative sense of ourselves".<sup>104</sup> This correspondence should be understood less strictly than in the other hierarchical accounts described above: Christman aims at providing a criterion of authentic acceptance of values and character traits which is less demanding than Dworkin's 'identification' or Frankfurt's 'endorsement'.<sup>105</sup> According to Christman, this makes his model more open to the "ambivalences" and "inner tensions" which are part of most agents' lives.<sup>106</sup> Christman's solution is the introduction of a 'non-alienation' criterion, which he describes as follows:

---

<sup>103</sup> Ibid.

<sup>104</sup> Ibid., 146.

<sup>105</sup> See footnote 70 (on the evolution of Dworkin's model) above.

<sup>106</sup> Ibid., 143.

If to be autonomous required that we *positively* value each aspect of our motivational structure and conditions of action, only the supremely lucky and fulfilled among us would count as autonomous. The rest of us sorry sorts would have to make do with imperfections, ambivalences, and the acceptance of tragedy in our lives. In order to keep “autonomy” from simply being a label for the ideal life [...], we should require of the autonomous person only that she accept herself *in the minimal sense* of not being *acutely alienated* from the basic elements in her motivational structure and life conditions.<sup>107</sup>

Therefore, autonomy does not require that we actively and emphatically endorse aspects of ourselves but that we at least succeed in not rejecting them. Indeed, according to Christman, we can come to terms with aspects of ourselves which are “not ideal from our point of view, but also which are not compulsions to which we are resistant”.<sup>108</sup>

In Christman’s view, being alienated from a value implies both *cognitively and affectively* resisting this value: Christman defines alienation as “a combination of judgement and affective reaction”.<sup>109</sup> When one is alienated from a particular trait, in addition to *judging* that a value or character trait is not compatible with who one is, one will *experience* an “anxious sense that the factor in question is constraining, that it undercuts one’s settled motivational frame and sense of the validity of that frame (given surrounding conditions).”<sup>110</sup>

---

<sup>107</sup> Ibid., 13; emphases added.

<sup>108</sup> Ibid., 143.

<sup>109</sup> Ibid., 144. As I explain in Chapter 3, in more recent versions of Christman’s model, it is suggested that the affective component of the experience may be sufficient evidence of the underlying judgement that alienation implies. Therefore, the affective component might be enough to identify alienation.

<sup>110</sup> Ibid.

Second, Christman also speaks about ‘authentic reflection’.<sup>111</sup> This second sense of authenticity requires “authenticity checks” that concern the reflective procedure itself and guarantee that critical self-assessment is *performed* authentically. According to Christman, one’s basic commitments are *also* part of “the *executive* function by virtue of which reflective judgment is made”.<sup>112</sup> For example, assuming minimal rationality, an agent could have a tendency to be more or less thorough in her self-assessments – i.e. she could consistently fall within more “impulsive” or “over-thinking” patterns – when evaluating her character traits. In practice, authentic reflection is secured through a ‘sustained critical reflection’ condition, defined as “[c]ritical self-reflection repeated in a variety of contexts with similar evaluative results.”<sup>113</sup>

Christman’s revised model can be schematized as follows: An agent is autonomous vis-à-vis an organising value if: a) she is *competent* for reflection; and b) should she reflect on an organizing value she would do so both critically and according to an *authentic pattern of reflection*; and c) should she reflect on an organizing value, she would *not* feel deeply *alienated* from it given both the history of the value and the agent’s narrative self-conception.

For the sake of clarity, let’s illustrate how this model works by considering one of Christman’s examples: A person with a history of childhood abuse could go through a process of revision of the memories of her abuse and deeply reject all the values and skills that were acquired as a result of her time with her abuser. Her self-revision would count as autonomous as long as she reasons in a way that is competent and can be sustained over time (she is not in an incontrollable rage, for example, and judges according to her sustained commitments, her personal style of reflecting, etc.). Christman also exemplifies what heteronomy would look like by imagining that this agent has (and enjoys) the skill of playing

---

<sup>111</sup> Ibid., 149.

<sup>112</sup> Ibid., 150; emphasis added.

<sup>113</sup> Ibid., 152-3.

the piano without remembering that her abuser is also responsible for her piano skills (let's say she was in fact taught to play the piano by the abuser). Then, Christman claims that she is not autonomous with respect to this character trait because, *were* she aware of the fact that her abuser was responsible for her piano skills, she would feel acutely alienated from them.<sup>114</sup>

Before concluding this chapter, I want to briefly consider the issue of the *scope* of the historical information that should inform one's self-assessment. References to one's "psychological history" or to one's "self-development" are not helpful to determine exactly what kind of historical materials Christman considers relevant information to assess one's autonomy. While an agent's biography (i.e. her past experiences) are certainly part of one's psychological history, other relevant intersubjective or social elements (e.g. family or cultural histories) are less obviously so. Still, Christman's explicit aim is to consider agents as *socio-historical* selves who are, for example, affected by the social meanings of the categories they use in self-definition.<sup>115</sup> How narrowly or broadly should we define the "personal histories" of socio-historical selves?

In his descriptions and in his examples Christman seems to oscillate between purely biographical accounts of one's history (like in the piano example presented above) and a more "fundamental" social history. Indeed, in some of his later writings his focus seems to have changed from personal history to something broader: the sociohistorical conditions of possibility of one's character and values.<sup>116</sup>

The same move towards a "broader" conception of one's history (i.e. one that includes not merely personal elements but also social ones) could be signalled when Christman considers potential "sources" of historical information. One obvious source is memory and

---

<sup>114</sup> Ibid., 158.

<sup>115</sup> Ibid., 209.

<sup>116</sup> John Christman, "Autonomy and Liberalism: A Troubled Marriage?," in *The Cambridge Companion to Liberalism*, ed. Steven Wall (Cambridge: Cambridge University Press, 2015), 151.

Christman gives a quite predominant role to “the powers of memory” in his theory – thanks to memory one would gain access to one’s “ongoing historical narrative”.<sup>117</sup> Prima facie this focus on memory may appear to restrict the relevant historical materials for autonomy to those contained in the collection of facts that constitute one’s biography, since one does not strictly “remember”, for example, social constraints.

But what about other relevant influences on one’s personal history that one could not learn about through “the powers of memory”? Think for example of the way in which (past) conditionings may include elements that one could not actually remember, like (say) relevant ancestors one did not get to meet or family migrations one did not live. I believe that it is quite uncontroversial to claim that these kind of family and/or social conditionings could shed relevant light on one’s character or dispositions.<sup>118</sup>

Christman’s account of memory seems to allow for the consideration of memories that are not personal, he claims:

[...] memory is crucial to agency, as we saw. Therefore, persons must be modelled as concerned with not only their current status and future prospects but also the way in which they have access to and can understand their memories, both *public and private*. This means that in order to grasp the meaning and importance of one’s most fundamental experiences – the constituent element of one’s autobiographical narrative – one must reflect on those events by way of concepts and symbols in the public discourse.<sup>119</sup>

---

<sup>117</sup> Christman, *The Politics of Persons*, 86

<sup>118</sup> There is even a growing literature (mainly in psychoanalysis) on how trauma is passed on from one generation to the next (often two generations down from those who experienced it first-hand). See for example Jill Salberg, “The Texture of Traumatic Attachment: Presence and Ghostly Absence in Transgenerational Transmission,” *The Psychoanalytic Quarterly* 84, no. 1 (2015).

<sup>119</sup> Christman, *The Politics of Persons*, 112; emphasis added.

This passage is interesting for many reasons. First, because accounting for the existence of a “public” memory allows for the integration of non-personal histories in Christman’s account: if we gain relevant historical information both through public and private memories, then relevant historical materials should not be restricted to the past experiences of individual agents. Furthermore, acknowledging the fact that one makes sense of one’s ongoing historical narrative thanks to concepts and symbols that are part of the “public discourse” makes even the “personal” historical materials that one can find in one’s biography already mediated by social history. After all, possible *readings* of one’s history and the identification of (positive and negative) relevant events will be affected (e.g. enabled or not) by one’s context (e.g. by available concepts, as I explain in Chapter 3).

This is, I believe, the kind of framing of the historical conditions in a model of autonomy which is helpful when one aims at considering self-government and oppression together. I believe that it is quite uncontroversial to claim that oppressive interferences have a history that goes *beyond* individuals’ psychological histories. Therefore, making sure that agents possess, when reflecting, not only information about their biographies but also, more broadly, about relevant social influences or conditionings seems exactly right and necessary when one is dealing with oppression – i.e. with a phenomenon that might manifest in an individual’s life but which has a more structural or social nature.

Moreover, if we acknowledge that, even when we extract historical materials from our personal histories, we are performing an interpretative task that is socially mediated, then we need to consider how relevant it is for self-assessment to reflect on available social meanings. As I argue in the following chapters, if awareness of one’s history is to have any emancipatory potential whatsoever, this cannot happen by leaving unproblematized the way in which one structures and reads one’s history. Indeed, historical theorists of autonomy should reflect more than they do on the risks of taking “dangerous” interpretative paths vis-à-vis one’s

history. Crucially, it is necessary to say more on the measures taken to avoid readings of one's history that could actually contribute to one's oppression.

In spite of the need to refine Christman's historical condition, I believe that the inclusion of this condition in his model is a key contribution to theories of autonomy and constitutes a further reason why Christman's account is a promising starting point for my project. Let me just be clear that my assessment of the relevancy of Christman's historical model relies on the assumption that "personal history" should *not* be restricted to personal biographical history and, even, on the stronger assumption that biographical history *alone* is insufficient. In Chapters 2 and 3, I further disambiguate "personal history" for the purposes of my project and, more importantly, I also provide a *rationale* for interpreting the historical condition within a historical model of personal autonomy both in a broader sense (i.e. as social history) and in a more fundamental way (i.e. as the historical conditions of possibility of certain values or character traits).



## Chapter 2: Theoretical Resources for a Critique of Historical- Procedural Models

My aim in this chapter is to provide some conceptual tools from the philosophy of Michel Foucault that could help strengthen the historical-procedural models presented in the previous chapter. I have claimed that historical accounts (and Christman's model in particular) constitute a promising starting point for my project but, at the same time, I have gestured towards some vagueness and areas of tension that require clarification. In this chapter, I connect these proto-critiques with conceptual tools that I use, in the chapters that follow, for a more developed critique (Chapter 3) and, eventually, for a proposal that overcomes these difficulties (Chapter 5).

I proceed in the following way: I present three main areas of tension that emerge from my previous presentation of historical accounts. I then connect these areas of tension with relevant Foucauldian considerations which will help both to justify why one should consider these issues problematic and to envisage, in following chapters, a way of overcoming these insufficiencies.

The key areas of tension within historical-procedural models that I expand on in this chapter are: 1) the somewhat incomplete account of 'procedural independence'; 2) the vagueness vis-à-vis the historical materials that are considered relevant and sufficient to perform a critical self-assessment; and 3) the limited understanding of oppression that we would need to presuppose to deal with oppression through available procedural models of autonomy. Let me briefly explain what is under analysis in each case.

1) **The somewhat incomplete account of 'procedural independence'**. As defined earlier, this notion identifies both subjective and intersubjective conditions of possibility for reflective self-assessment. Namely, externalist theorists aim at ensuring both *non-coerced* and *competent* reflection, and

*critical* reflection: proceduralists do not confine themselves to securing non-manipulated and minimally rational reflection, but they also expect that agents do not carry out self-assessments by “blindly” resorting to any value-framework that they might find themselves with. Historical conditions are typically added to procedural models to avoid some forms of blind acceptance of value-frameworks – e.g. acceptance of a value-framework should at least take into account the *origin* of the value-framework. While this historical move is promising, is it enough to secure a *critical* perspective?

I turn to the Foucauldian notion of ‘critical ethos’ to analyse how critical self-assessment might be enhanced through a special relation to one’s history or to what is historical in oneself. Crucially, seeing critical self-assessment as an “ethos” or “attitude” has the advantage of not reducing procedural independence to a condition of possibility for trustworthy critical reflection which, when externally absent, blocks autonomy – i.e. in certain cases that I define later, agents can “work” on their procedural independence and *enhance* their “independence of mind”. Moreover, as I argue below, this critical attitude is also fruitful to deal with “interferences” non-reducible to forms of manipulation or deception.

Finally, I problematize the notion of ‘reflective competence’ in light of this (redefined) critical attitude. I do so by connecting the definitions of reflective competence available in procedural models with Foucault’s early writings on madness and unreason and with the feminist reception of Foucault.

- 2) **The vagueness vis-à-vis the historical materials that are considered relevant and sufficient to perform a critical self-assessment.** I have suggested that historical materials are underdetermined for two reasons:

First, it is not sufficiently clear what is englobed by the term “psychological history” in historical accounts and, as I have suggested, this vagueness or ambiguity may reduce the effectiveness of a procedural test in some oppressive contexts. Moreover, it is not merely necessary to disambiguate the notion of “psychological history”, but we also need to provide a *justification* for opting for different kinds of history (e.g. biographical, familial, social) in the context of a project that seeks to secure *critical* self-assessment in contexts of oppression.

Second, I have claimed that the *interpretative* nature of one’s history is not sufficiently problematized. For example, historical procedural accounts which aim at unveiling the effects of oppression on one’s character should avoid narratives where the origins of one’s character are (biographically) explained but characters are (socially) unproblematized. Moreover, since establishing the conditions of possibility of experiencing alienation vis-à-vis one’s character or values is key, the different affective effects of different ways of telling histories should be considered.

I turn to the notion of ‘genealogy’ as a way of dealing with these two issues.

3) **The limited understanding of oppression that we would need to presuppose to deal with oppression through available procedural models of autonomy.** As I explain below, historical accounts are not sensitive enough to detect heteronomous interferences which do not happen in a “disruptive” way but, rather, affect our developments as a whole. I unpack and justify this claim thanks to Foucault’s analysis of power relations. Crucially, I look into what differentiates power relations of different kinds from ‘states of domination’ and ‘relationships of violence’.

In what follows, I consider these issues in turn and present the abovementioned Foucauldian tools.

### I. What Counts as ‘Procedural Independence’?

The idea that reflective self-assessment needs to take place under certain conditions for it to be “genuine” appears in the three externalist-procedural models that I have analysed in the previous chapter in different (though closely connected) ways. Amongst these conditions, we could draw a distinction between *intersubjective* and *subjective* conditions.

By ‘intersubjective’ conditions, I refer to the conditions that procedural models typically set in place to secure reflective independence from others. As I have explained before, cases in which one is subjected to (and therefore not independent from) a manipulator’s or deceiver’s will (i.e. an *alien* will) would be typically incompatible with procedural independence.<sup>120</sup>

By ‘subjective’ conditions, I refer to those “mental states” according to which available models typically determine whether agents are capable of carrying out an adequate reflective assessment of themselves and their conditions. In order to establish that an agent meets the subjective requirements to perform a reflective self-assessment, it is necessary that both: (i) certain *negative* mental states are *ruled out*; and (ii) the possession of certain *capacities* is *verified*. Concerning (i), an agent is in a position of procedural independence when she is not in a mental state that could cast doubts on the validity of her reflection or the results of her reflection. For example, being furious, on drugs or hallucinating are typically considered as instances of such untrustworthy scenarios.<sup>121</sup> Concerning (ii), an agent who is a suitable

---

<sup>120</sup> See Dworkin, “Autonomy and Behaviour Control,” 25; Christman, *The Politics of Persons*, 146; Mele, *Autonomous Agents*, 148-9.

<sup>121</sup> Christman, *The Politics of Persons*, 147.

candidate for autonomy needs to be reflectively competent. The latter is the case when, for example, an agent is “minimally rational”.<sup>122</sup>

Based on my analysis in Chapter 1, we can signal some general tendencies in proceduralists’ definitions of ‘procedural independence’:

First, *intersubjective* procedural independence is mostly defined by procedural theorists in a negative fashion – i.e. as the *absence* of (past or present) illegitimate *interferences*. When interferences are not found (and one is competent for reflection) then agents’ value frameworks are typically seen as trustworthy enough to carry out a procedural self-assessment. As I explained in Chapter 1, manipulation and deception are the paradigmatic cases of heteronomous interference which violate procedural independence. Putting forward manipulation and deception as the main models of heteronomous interference has consequences when it comes to imagining the logic and temporality of possible interferences. For example, manipulation and deception typically suggest a dyadic form – i.e. one agent is manipulated/deceived by another, generally identifiable, agent(s). Moreover, manipulation and deception typically suggest that the one who manipulates or deceives acts *intentionally* and not, for example, accidentally or unwillingly.

I want to argue that seeing heteronomous interferences in the way described above, may have unfortunate consequences given the purposes of this thesis. Namely, picturing heteronomous interferences as typically dyadic and subjectively intentional (i.e. as typically in the hands of agents who intend to manipulate or deceive) may make it difficult for a procedural model to accommodate oppression in all its forms. Indeed, we can even say that trying to fit possible oppressive interferences into the schema presupposed by relations of manipulation and deception would imply adopting a ‘neoliberal’ understanding of oppression, in the sense proposed by Ladelle McWhorter. According to McWhorter, in a

---

<sup>122</sup> Dworkin, “Autonomy and Behavior Control,” 27; Christman, *The Politics of Persons*, 152-3; Mele, *Autonomous Agents*, 187. The three positions are developed and analysed in more detail in section *b* below.

neoliberal sense derived from Hayek, oppression is equated to a “state of continuous acts of coercion” and the latter implies “that I still choose but that my mind is made someone else’s tool.”<sup>123</sup> McWhorter argues that the latter has dangerous implications because oppression is made “a matter of two or more people’s mental states—that of conscious intentionality on the part of the oppressor and that of appropriated volition on the part of the oppressed—not of social or political circumstances.”<sup>124</sup>

The case of gender oppression is a useful example here. If we accept that being socialized as a ‘woman’ in a patriarchal context makes women vulnerable to gender oppression, then we have a case that defies the picture of oppression sketched above. Even if gender oppression could, in a sense, be considered the result of the interventions of individual others (e.g. one’s caregivers, one’s partners, etc.), it is certainly not reducible to those interventions. Think for example how quitting a particularly sexist personal relationship may improve one’s life conditions in many ways but does not suppress one’s oppression *qua* woman. Crucially, gender oppression does not require “conscious intentionality” on the part of those who perpetuate it or enforce it – indeed, as feminists have noted, gender oppression can be reproduced in spite of everyone’s good intentions and even within relations of love and care.<sup>125</sup>

To be sure, externalist procedural theorists *do* provide some tools to integrate social or political circumstances into their accounts. (Indeed, I have hinted already at some ways in which Christman’s historical conditions can accommodate social and political conditionings. I say more about this in Chapter 3.) My point for now is, however, that whatever space we make to consider oppression in procedural terms will be seriously flawed if our warning signs

---

<sup>123</sup> Hayek in McWhorter, “Post-Liberation Feminism and Practices of Freedom,” 64.

<sup>124</sup> Ibid.

<sup>125</sup> See footnote 21 above, on Allen’s example.

are mainly triggered by forms of coercion like the ones described by McWhorter above, namely by *intentional* acts of coercion carried out by *concrete* oppressors.

A second set of difficulties emerges when we look into the subjective conditions of possibility for reflective (self)assessment – i.e. when we consider ‘competence’ conditions. Definitions of “rationality” are not exactly innocent, and procedural accounts need to explicitly consider which normative commitments are imported into their models through their competence conditions *in spite of* their professed content-neutrality.

I suggest that different sets of Foucauldian tools could help shed light on the abovementioned problems. First, Foucault’s reflections on the ‘critical ethos’ are relevant to tackle those issues which arise when we try to define intersubjective conditions for procedural independence. Foucault’s critical attitude is motivated by reasons that are close to the ones that motivate historical theorists: in both cases they seek to promote (history-sensitive) forms of self-interrogation which could help to set limits to forms of “government” (as defined below) in a particular socio-historical configuration. In addition to this, I argue, Foucault’s critical attitude successfully deals with the insufficiencies signalled before: namely, it does not presuppose that possible interferences are necessarily disruptive or that they involve concrete others with intentions to manipulate, deceive or coerce.

Second, Foucault’s reflections on madness and unreason allow for a better understanding of the dangers associated with the reflective competence conditions integrated into procedural models. Crucially, what counts as “rational” in a specific sociohistorical situation is frequently shaped in light of social ideals, injunctions, and aims which cannot be so easily separated from power configurations. Ideals of, for example, economic productivity may end up influencing the way in which we define reflective competence and attribute it (or deny it) to agents. Even though, as I discuss below, complete neutrality may be incompatible with defining a model of autonomy, the abovementioned dangers deserve close attention in the context of a study of autonomy under oppression.

Let me now present the necessary theoretical elements to unpack and ground these claims.

a) (Re) defining ‘intersubjective’ procedural independence:

Foucault repeatedly considered the question of ‘critique’ in relation to Kant’s answer to the question *Was ist Aufklärung?*<sup>126</sup> Foucault’s reflections on critique are tightly connected with his historical analysis of ‘government’ and ‘governmentality’. This relation between the “problem” of government and the “attitude” of critique is crucial because, as I show below, critique needs to be understood precisely as the possibility of *limiting* government.<sup>127</sup> Let me unpack this claim by briefly introducing the key elements at stake:

First, it is necessary to say a bit more about government and, more specifically, about the type of government which, according to Foucault, is limited by critique. Foucault signals an “explosion” of government “as a general problem” in Europe after the sixteenth century.<sup>128</sup> The “problem” of government includes matters as broad as: “[h]ow to govern oneself, how to be governed, how to govern others, by whom the people will accept being governed, how to become the best possible governor...”.<sup>129</sup> In sum, “government” should

---

<sup>126</sup> Foucault discusses the text published in *Berlinische Monatsschrift*, Dec. 1784, vol. IV, 481-491. In a 1978 lecture at the *Société Française de Philosophie*, Foucault gives a talk titled *What is Critique?* but admits that he hesitated on whether to title his talk *What is Aufklärung?* By virtue of its contents, this talk can be grouped with other texts in which Foucault explores the questions of modernity and the critical attitude. I refer also to two texts with the same title – i.e. *What is Aufklärung?* – published in 1984, one in the US, and the second (an extract from a January session of his 1983 course at the Collège de France) in France. Finally, in the beginning of a conference (*La Culture de Soi*) given by Foucault in Berkeley in April 1983, he reanalyses Kant’s text on Enlightenment. The strategies and emphasis of the different texts varies. For an analysis of the main shifts between Foucault’s approach in 1978 and Foucault’s way of dealing with the critical attitude during Berkeley’s lecture of 1983, see Daniel Lorenzini and Arnold Davidson, “Introduction,” in Michel Foucault, *Qu’est-ce que la Critique ? suivi de La Culture de Soi*, (Paris: Vrin, 2015), 11-30.

<sup>127</sup> By ‘attitude’, Foucault refers to “a certain way of thinking, speaking and acting, a certain relationship to what exists, to what one knows, to what one does, a relationship to society, to culture and also a relationship to others...” (Foucault, “What is Critique?,” 24). Foucault also uses the term “virtue” and, in the French manuscript, he even refers to ‘critique’ as a ‘style’. (Foucault, “Qu’est-ce que la Critique?,” 35) Moreover, the critical attitude or *ethos*, Foucault warns, should not be mistaken for “a theory, a doctrine, nor even as a permanent body of knowledge that is accumulating.” (Foucault, “What is Enlightenment?,” 319). In sum, it should be clear that a ‘critical attitude’ is not reducible to an “intellectual” critique. I come back to this below.

<sup>128</sup> Michel Foucault, “Governmentality,” in *Power Vol. 3 of Essential Works of Michel Foucault 1954-1984*, ed. James D. Faubion (New York: The New York Press, 2000), 201.

<sup>129</sup> *Ibid.*, 202.



not be understood in an exclusively juridical sense (e.g. as the issue of how a state should govern legitimately).

This historical “explosion” of government which Foucault refers to is also linked to the expansion, to secular contexts, of a particular form of government called ‘pastoral power’. As it is well known, Foucault traces the notion of pastorship back to the image of the shepherd looking after and guiding his flock, present in ancient religious Hebraic texts and, with differences, in early Christian texts.<sup>130</sup> Western European societies, Foucault claims, “evolved a strange technology of power treating the vast majority of men as flock with a few as shepherds.”<sup>131</sup>

Crucially, the techniques of government derived from pastorship aim at governing individuals in a “continuous and permanent way”.<sup>132</sup> To achieve this, they create, study, and target specific ‘populations’. ‘Populations’ are groups of living individuals of the same “species” (e.g. individuals who share biological features) which become objects of a knowledge aimed at managing these individuals *en masse*. The latter, however, should not obscure the fact that pastoral technologies are also concerned with the life of *each and every one* of the members of a certain population.<sup>133</sup> Moreover, as I explain in section III below, the governmental technologies derived from pastorship concern each individual in an external manner (as bodies) but *also* “internally” – i.e. they concern what in religious terms could be presented as one’s “conscience” or “soul”.<sup>134</sup> In non-religious terms this “internal” dimension could be introduced more broadly as “the problem of individuality” or of “self-identity”.<sup>135</sup>

---

<sup>130</sup> Michel Foucault, “Omnes et Singulatim: Toward a Critique of Political Reason,” in *Power Vol. 3 of Essential Works of Michel Foucault 1954-1984*, ed. James D. Faubion (New York: The New York Press, 2000), 300-11.

<sup>131</sup> *Ibid.*, 303.

<sup>132</sup> *Ibid.*, 300.

<sup>133</sup> *Ibid.*, 309-310.

<sup>134</sup> *Ibid.*, 310.

<sup>135</sup> *Ibid.*, 300.

Another way of introducing this process of ‘governmentalisation’ of the modern state within Foucault’s work is to present it as the integration of two modern technologies of power: an “individualising” one, i.e. ‘discipline’; and a “totalizing” one, i.e. ‘biopower’.<sup>136</sup> As I briefly mentioned in the Introduction, according to Foucault disciplinary power was “invented” in the seventeenth and early eighteenth centuries, and it is a form of power which targets primarily individuals’ bodies and what these bodies do.<sup>137</sup> Crucially, disciplines create discourses, “knowledges and multiple fields of expertise”, which set “codes of normalization” in light of which individuals are permanently monitored, examined, and valued.<sup>138</sup> Regarding biopower, Foucault defines it as a form of power which emerged later in the eighteenth century and which targets the “man-as-living-being” or the “man-as-species”, i.e. it targets ‘populations’ (as defined above).<sup>139</sup> Biopower aims at regularizing the life-processes of these populations by, for example, intervening to “maintain averages” or “compensate for variations” in birth or mortality rates.<sup>140</sup>

Amidst this multiplication of questions about government, of realms of possible government, and of techniques of government, Foucault claims, critique asks the question of government *negatively*, namely as “how *not* to be governed?”<sup>141</sup> Foucault claims:

Facing [governmentalisation] head on and as compensation, or rather, as both partner and adversary to the arts of governing, as an act of defiance, as a challenge, as a way of limiting these arts of

---

<sup>136</sup> As Lois McNay puts it: “Governmental reason represents an approach to social control that operates not through direct state sanction but through the indirect shaping of ‘free’ social practices on two levels: *regulatory or massification techniques* that focus on the large-scale management of populations (for example, controlling mortality levels and birth rates, planning of urban environments or investing in ‘human capital’) and *individualizing, disciplinary mechanisms* that shape the behaviours and identity of the individual through the imposition of certain normalizing technologies or practices of the self.” (McNay, “Self as Enterprise,” 57). See also Allen, *The Politics of Our Selves*, 56.

<sup>137</sup> Foucault, *Society Must be Defended*, 35.

<sup>138</sup> *Ibid.*, 38.

<sup>139</sup> *Ibid.*, 242.

<sup>140</sup> *Ibid.*, 246-7.

<sup>141</sup> Foucault, “What is Critique?,” 28, emphasis added.

governing and sizing them up, transforming them, of finding a way to escape from them or, in any case, a way to displace them, with a basic distrust, but also and by the same token, as a line of development of the arts of governing, there would have been something born in Europe at that time, a kind of general cultural form, both a political and moral attitude, a way of thinking, etc. and which I would very simply call the art of not being governed or better, the art of not being governed like that and at that cost. I would therefore propose, as a very first definition of critique, this general characterization: the art of not being governed *quite so much* [*l'art de n'être pas tellement gouverné*].<sup>142</sup>

Foucault warns that this critical question cannot be asked or understood in absolute terms – i.e. critique does not aim at a situation in which one is not governed *at all*. Rather, this counter-governmental attitude is always asked “locally” – i.e. as the possibility of not being governed “like that,” “by them”, “for that”.<sup>143</sup> However, I argue, this “local” character of critique should not be exaggerated either insofar as critique does not seem reducible to a matter of choosing between different forms of, say, equally bad, equally all-encompassing, or equally subordinating government. Foucault can be read (and translated) to say here that it is possible to *limit* government. Indeed, I want to suggest that, thanks to critique, it is possible to become *less* governed by others, by certain power configurations, and even (as I explain below) by certain “truths” about oneself.

---

<sup>142</sup> Ibid., 28-9; Foucault, “Qu’est-ce que la Critique ?,” 37. As I commented in footnote 38 in the Introduction, the French phrase in italics can be read as indicating more than a choice between different forms of being governed, namely as implying the possibility of being governed *less*. I follow the translator here to read it in the second sense.

<sup>143</sup> Ibid.

Admittedly, we need to ask *how* this critique is possible. Foucault provides some historical examples of this critical attitude. With some variations, these examples describe instances where an authority is questioned and this “authority” should be understood broadly, that is, not necessarily as a specific person (e.g. a sovereign) but also as a religious tradition, a dogma, a corpus of laws.<sup>144</sup> Crucially, in these examples, Foucault links the “art of critique” with the task of undoing the “bundles” of power, truths, and subjects, thanks to which government is possible.<sup>145</sup> That is, one of the aims of critique is to de-subject individuals who are “caught” within a certain “politics of truth” – and by “caught” I mean not merely that they, as individuals, play a role within this configuration of truth and power, but also that they, subjects, are an *effect* of these “truths” imposed on them.<sup>146</sup> As I have already explained in the Introduction (and as it becomes clear below), this “effect” should neither be understood as an absolute determination nor as something that merely “happens” to subjects: individuals in a sense “*self-fashion*” in light of a “truth” and, therefore, this process of self-formation according to “truths” is not completely beyond the reach and control of individuals.

In order to explain a bit more how Foucault sees the undoing of “bundles” of power, truths, and subjects as possible, it is necessary to turn to those texts in which Foucault defines critique in more direct relation with Kant: Foucault is interested in a form of critique found in Kant’s reflections on Enlightenment and revolution which questions ‘present reality’ (or contemporary reality).<sup>147</sup> Foucault claims:

---

<sup>144</sup> Foucault, “What is Critique?,” 30-1.

<sup>145</sup> Ibid., 32.

<sup>146</sup> Ibid., 32

<sup>147</sup> Foucault’s references to Kant’s reflections on revolution could be considered as another attempt to illustrate what discontinuity in the present, unpredictability, and a special enthusiastic attitude towards the future look like: “it seems to me that these two questions – “What is Aufklärung?” and “What is the Revolution?” – [...] are the two forms in which Kant poses the question of his own present reality” (Michel Foucault, *The Government of Self and Others, lectures at the Collège de France 1982-1983* (Hampshire: Palgrave Macmillan, 2010), 20).

This other critical tradition does not pose the question of the conditions of possibility of a true knowledge; it asks the question: What is present reality [*actualité*]? What is the present field of our experiences? Here it is not a question of the analytic of truth but involves what could be called an ontology of the present, of present reality, an ontology of modernity, an ontology of ourselves.<sup>148</sup>

Then, interrogating one's present reality, in the way Kant does when he reflects on Enlightenment or revolution, is a way of looking for the present (historical) conditions of possibility of our experiences.

The Kantian association should be taken with caution insofar as Foucault is not interested in the "field" or limits of our experience in the same way that Kant's transcendental philosophy is. For Foucault, the limits of present reality should be found to later *interrogate its necessity* and, therefore, to open up the possibility of *transgressing* them. Foucault explains:

...if the Kantian question was that of knowing [*savoir*] what limits knowledge [*connaissance*] must renounce exceeding, it seems to me that the critical question today must be turned back into a positive one: In what is given to us as universal, necessary, obligatory, what place is occupied by whatever is singular, contingent, and the product of arbitrary constraints? The point, in brief, is to transform the critique conducted in the form of necessary limitation into a practical critique that takes the form of a possible crossing-over [*franchissement*].<sup>149</sup>

---

<sup>148</sup> Ibid., pp. 20-21.

<sup>149</sup> Foucault, "What is Enlightenment?," 315.

So, the critique that Foucault refers to is different from transcendental philosophy because it does not renounce going *beyond* what is given to us as a fixed condition of possibility of our experiences.

But what does exactly “going beyond” limits mean here? On the one hand, critique seems to be likened to a “modal” intellectual change: what we thought *necessary* is now seen as *contingent* and, therefore, we realize that things *could* be otherwise. However, the use of terms like “attitude” or “art”, and the characterisation of critique as a “practical” matter, also suggest that critique is not reducible to an intellectual task. Indeed, Foucault claims, “the critique of what we are is at one and the same time the historical analysis of the limits imposed on us and an experiment [*épreuve*] with the possibility of going beyond them [*de leur franchissement possible*].<sup>150</sup> In other words, actual practices of “testing” limits, should also be considered critical in the sense proposed by Foucault.

As I suggested earlier, this interrogation of present reality has effects on our *self*-understanding. This “modal change” explained above also applies when it comes to interrogate the way in which we see ourselves. For example, Foucault sees in this Kantian interrogation of his present not merely an attempt to unveil a tradition which must be continued or a universal anthropological truth about who ‘we’ are.<sup>151</sup> Foucault notes that Kant reflects on his belonging to a “*particular ‘we’ [...] linked, to a greater or lesser extent, to a cultural ensemble characteristic of his contemporary reality [actualité]*.”<sup>152</sup> In other words, critique is a “historical investigation into the events that have led us to constitute ourselves and to recognize ourselves as subjects of what we are doing, thinking, saying.”<sup>153</sup> Like above, unveiling the historical determinations which made “us” (qua subjects of a particular time in

---

<sup>150</sup> Ibid., 319.

<sup>151</sup> Foucault, *The Government of Self and Others*, 13.

<sup>152</sup> Ibid.; emphasis added.

<sup>153</sup> Foucault, “What is Enlightenment?,” 315.

history) what we are, also entails “the possibility of no longer being, doing, or thinking what we are, do, or think.”<sup>154</sup>

A key point here is that, when it comes to carrying out a critique of our historical mode of being, the possibility of being otherwise is not merely opened up as a result of unveiling a, so to speak, passive historical contingency (e.g. by pointing out that we are the “products of our time”). Additionally, Foucault explains, when we turn to our “ontological history” (as he puts it in his 1983 Berkeley lecture), we gain insight into the ways in which *we* formed *ourselves* through history, e.g. by self-relating how, by “fashioning” these self-relations through which techniques.<sup>155</sup> At this occasion, then, Foucault explicitly connects the critical attitude with his *ethical* studies on sexuality (about which I say more below). Therefore, what should be noted is that critique opens up possibilities of *being* otherwise not exclusively by challenging “necessary truths” about ourselves, but also by showing that being one way or another is, to a certain extent, in our hands.

To recapitulate the points made so far, I have shown that interrogating one’s present reality entails a special questioning of *ourselves*. Therefore, from what we have seen up to now, if there is a form of self-questioning which is enabled by the critical attitude, this is one which happens at the level of a “we”. That is: questioning “oneself”, in this context, implies questioning a form of being which is not exclusively individual but also that of one’s contemporaries. Furthermore, the interrogation of our present reality and of ourselves does not require that the agent who carries out critique *continues* a previous tradition or remains “faithful” to a historical mode of being. A historical critique of ourselves connects the subject who interrogates herself with her sociohistorical situation, but from this it does not follow that history allows to unveil something necessary or unavoidable about one’s way of being.

---

<sup>154</sup> Ibid., 315-6.

<sup>155</sup> Michel Foucault, “La Culture de Soi,” in Foucault, *Qu’est-ce que la Critique ? suivi de La Culture de Soi*, 84.

Indeed, historical determinations are framed as *contingencies*, as “inherited” modes of being that *could* be changed by those affected.

To conclude, let me address more explicitly why I consider that a Foucauldian notion of critique deals better with the issues signalled in the beginning of this section:

First, as it emerges from the characterisation of critique above (and as my analysis below will further develop), a Foucauldian approach could avoid the reductive understanding of ‘interference’ that is presupposed by procedural theories of autonomy. Within a Foucauldian framework there is space to reflect on (and to limit) historically-conditioned ways of being which are not necessarily implemented through dyadic disruptive interventions on individuals’ lives. Crucially, even settled forms of self-relating and those modes of being that one considers more “natural” might open the door to a higher degree of government by others (i.e. to a higher degree of heteronomy). In a word, sustained influences and settled ways of being are, in this context, at least as dangerous as obvious heteronomous interventions (e.g. manipulation or deception).

A further advantage of the Foucauldian framework presented above is that it calls into question whether heteronomous interferences are necessarily the product of intentional subjective interventions (which, as explained earlier, seems typically the case in the two paradigmatic models of interference proposed by proceduralists – i.e. deception and manipulation). As Foucault’s analysis shows, potentially constraining sociohistorical conditionings do have a logic and an aim, but they do not necessarily allow for the identification of particular others with intentions or wills that override those of the affected agent. Indeed, Foucault claims that “the logic of power” is “intentional and nonsubjective” – i.e. it has a clear direction or purpose but it is also “anonymous” (i.e. no one necessarily invented or formulated its aims).<sup>156</sup>

---

<sup>156</sup> Michel Foucault, *An Introduction, Vol. 1 of the History of Sexuality*, trans. Robert Hurley (New York: Pantheon Books, 1978), 94-5.



This way of thinking is, I believe, much more helpful to think about oppression than the schema provided by deception or manipulation. Consider, for example, the way in which the dynamics and structure of a racist society end up putting certain groups of individuals (e.g. white men) in a generally privileged position when it comes to influence or (even) direct others' decisions, conducts, value frameworks, etc. It would be inaccurate to claim, I argue, that the power of privileged individuals necessarily derives from them possessing a certain kind of knowledge – indeed, white privilege works even without individuals' awareness or understanding of structural racism. Moreover, it is not excessive to claim that in a racist society a white individual could be in a position of power regardless of their intentions. To be sure, I am not denying the existence of a structural “intention” to maintain some people in a position of power, my point is simply that at the level of individual interactions, we miss a big part of what is going on if we describe a structural problem as one of deception or manipulation.

To conclude this section, let me briefly go back to a point I made in the introduction to this chapter on the utility of a Foucauldian framework to think of reflective independence from pernicious or undesirable interferences not merely as an ideal or suitable starting point for critical reflection but also as that towards which critical reflection aims. In the understanding of critique presented above, critique seems to be not merely what one is able to achieve in the (right) reflective conditions but also the attitude of *actively seeking* to limit dependence on (certain) others and on (certain) “truths” about ourselves imposed on us.

Crucially, as I show throughout my thesis, agents can “start” self-critique by using those materials with which they find themselves, i.e. agents might resist or limit heteronomous influences by working on and with power-affected materials. For example, as I argue in Chapter 3, agents may become less governed by others without the need to access an ultimately “authentic” or independent self that could work as an Archimedean point to ground critique.

I restrict my claims above to the set of reflective conditions presented as ‘intersubjective’ because, admittedly, other rules apply to the set of conditions that define an agent’s reflective competence: it is a condition of possibility of any form of critical assessment to be minimally competent for reflection. I say more about competence conditions in the following section.

b) Problematizing ‘subjective’ procedural independence (i.e. ‘competence conditions’):

So far, I have provided resources to problematize those aspects of procedural independence that I labelled ‘intersubjective’. In this section, I deal with what I have called the ‘subjective’ dimension of procedural independence – namely, I turn to those conditions in procedural models meant to assess if an agent possesses and is able to exercise the necessary *capacities* to carry out any reflective self-assessment.

I proceed in the following way: First, for the sake of clarity, I briefly come back to the different characterisations of competence conditions available in the procedural models that I have discussed so far. Second, I turn to Foucault’s early reflections on madness and unreason and to the feminist reception of these reflections. My aim is to show that both the medical definitions of mental illness and the definitions of rationality typically embedded in competence conditions are not “neutral” but functional to social values within specific social configurations.

As I argued in Chapter 1, while procedural models aim at content-neutrality (i.e. they do not present specific values, preferences or lifestyles as being more or less conducive to autonomy), they still rely on normative criteria to define the procedures and necessary conditions which are constitutive of their models. For example, having “coherent” organising values (both synchronically –i.e. at a specific time in one’s life– and/or diachronically –i.e. throughout one’s life or development–) is generally deemed necessary for

autonomy in existing procedural accounts. While it is true that this condition may be satisfied by agents who endorse very different values, it is still contestable whether the general attitude of, for example, “remaining faithful” to one’s values throughout one’s life is preferable to taking a less-conservative approach towards one’s values or lifestyle. Indeed, a less-conservative approach might be preferable if one’s commitments have been shaped by the systemic values that one wishes to contest or, as I argue in Chapter 3, if one’s development as a whole has been affected by oppression.<sup>157</sup>

I argue that while getting rid of all values (e.g. epistemic or ethical) is probably incompatible with setting conditions for autonomy, if we want to secure opportunities for limiting oppression, we should at least avoid imposing on individuals those values and lifestyles functional to the oppressive contexts in which they live. In a word, being aware of the values present in (what Meyers calls) the ‘Constitutivity Axis’ of an account of autonomy is especially relevant in the context of a model of autonomy like the one I aim at in this thesis.<sup>158</sup>

Let me then start by briefly recalling the way in which Dworkin, Mele, and Christman define “reflective competence”:

For Dworkin, an autonomous agent needs to possess at least two important cognitive capacities: (i) “a second-order capacity [...] to reflect critically upon their first-order preferences, desires, wishes, and so forth”; and (ii) a capacity to accept or attempt to change these in light of higher-order preferences and values.<sup>159</sup> Even when these reflective capacities are not characterised in detail, the key idea that is conveyed is that autonomous agents need to be able to take critical *distance* from their first-order dispositions and ensure that they *cohere*

---

<sup>157</sup> Additionally, as I argued in Chapter 1, it should be noted that valuing “internal coherence” above other epistemic values may present the lives of those intersected by conflicting cultural injunctions as falling short of the necessary conditions for autonomy. This conclusion is contested by studies which put ‘intersectional identities’ in a privileged position for critical self-assessment. (See footnote 51 above)

<sup>158</sup> Meyers, “The Feminist Debate over Values in Autonomy Theory,” 115.

<sup>159</sup> Dworkin, *The Theory and Practice of Autonomy*, 20.

with higher-order dispositions. Dworkin also speaks more generally of the ability to “reflect rationally”, which should be preserved for the sake of autonomy by avoiding any methods of influencing people which could undermine such an ability.<sup>160</sup> All things considered, then, for Dworkin an autonomous agent should be capable of assessing particular contents in light of her values (i.e. higher-order dispositions) and she should be capable of doing so in a minimally rational manner.

Regarding Mele, he introduces three sufficient conditions for “psychological autonomy”.<sup>161</sup> Mele claims that (psychologically) autonomous agents: (i) are not compelled to possess a certain pro-attitude (e.g. a preference or desire); (ii) do not possess beliefs that are non-conducive to deliberation (i.e. beliefs that would compromise deliberation from an “executive” point of view); (iii) are reliable deliberators.<sup>162</sup> The latter involves possessing a set of “executive qualities” that could make critical assessment possible. ‘Quality’ here is a term that englobes “skills, capacities, and habits”.<sup>163</sup>

Crucially, Mele’s definition assumes that there are *objectively* “better and worse ways of deliberating”.<sup>164</sup> Standards of “deliberative excellence” are, in Mele’s words (and as they are in Dworkin’s model), defined from the point of view of “any *rational* deliberator”.<sup>165</sup> Indeed, this “objective” criterion of what constitutes reliable deliberation is so strong that any intervention on the individual that “improves” her executive qualities, even when the intervention itself bypasses control over one’s mental life, does not affect future autonomous

---

<sup>160</sup> Dworkin, “Autonomy and Behaviour Control,” 27.

<sup>161</sup> Mele claims that these conditions are “sufficient” and not “necessary and sufficient”, because he focuses here on a weaker, compatibilist, version of psychological autonomy which aims at defining what it means to have autonomous intentions. It is a “weak” version because it does not fully solve (or attempt to fully solve) the issue of whether intentions have been autonomously formed *if* one assumed a deterministic point of view. (Mele, *Autonomous agents*, 186) If we enter the compatibilist-determinist debate, Mele signals, it is extremely difficult to define what counts as “coercing” an intention (e.g. are all causally-determined intentions coerced?). Leaving this issue aside, Mele believes that these conditions are strong enough to rule out ‘covert non-constraining control’ – i.e. cases where a controller “covertly arranges ‘circumstances beforehand so that the agent wants and desires, and hence chooses and tries, only what the controller intends’.” (Kane in *Ibid.*, 179)

<sup>162</sup> *Ibid.*, 187.

<sup>163</sup> *Ibid.*, 182.

<sup>164</sup> *Ibid.*, 184.

<sup>165</sup> *Ibid.*

deliberation.<sup>166</sup> And what does this criterion of rationality include? Broadly, Mele associates rationality with “reliable” deliberation. In the case of instrumental rationality, for example, reliable cognitive procedures would typically improve the chances of achieving a certain aim and, conversely, “procedures driven by the anchoring effect or the confirmation bias” should be considered unreliable.<sup>167</sup>

Finally, when it comes to Christman, competence conditions require that the hypothetical reflection that is characteristic of autonomy is not “the product of social and psychological conditions that prevent adequate appraisal of oneself.”<sup>168</sup> He unpacks this claim by saying that this implies a “general capacity to *reflect adequately* without constriction, pathology, or manipulation.”<sup>169</sup> Additionally, as I have explained in Chapter 1, Christman makes room for the possibility of “personalizing” competent reflection so that it reflects the agent’s personality, her sustained mode of reflecting.<sup>170</sup> I deal with Christman’s model more extensively in Chapter 3, for now (and bearing in mind the more general problematisation that interests me here) this brief presentation of Christman’s competence conditions suffices.

From the definitions above (and excluding the more intersubjective aspects of these conditions, which have been already analysed above), we can extract three key elements that are part of the notion of reflective competence in externalist-procedural accounts: (i) “adequate” reflection (I take “adequate” to mean, “adequate given one’s social circumstances” and, even, “given the purposes of self-assessment”); (ii) the absence of mental pathologies; and (iii) minimal rationality (which typically entails being capable of

---

<sup>166</sup> Mele claims: “Suppose that there were a device that, when installed in any person’s head, would dispose the person to deliberate in ways that would reliably maximize his chances of locating efficient, effective means to his ends, whatever those ends might be. Imagine that neither the installation nor the operation of the device has any undesirable side-effects. [...] Rather, one’s capacity for deliberative excellence is enhanced. The installation and operation of such a device would not block autonomous deliberation—even though the very process of installation is one that bypasses the agent’s capacities for control over his mental life.” (Ibid., 183-4)

<sup>167</sup> Ibid., 184.

<sup>168</sup> Christman, *The Politics of Persons*, 146-7.

<sup>169</sup> Ibid.

<sup>170</sup> Ibid., 152-3.

logical thinking and having “coherent” value frameworks). Even if, at first glance, it might appear that the first feature is the only one which might let social normative commitments into a model of autonomy, I show below that the same might apply to the demands that agents are free from mental pathologies and that they are minimally rational. I justify these points by turning to Foucault’s analysis of madness and unreason.

In *History of Madness* Foucault challenges the view according to which the reforms in the treatment of the mentally ill that were put forward during the beginning of the nineteenth century represented progress towards a more “positivistic” treatment of mental illness.<sup>171</sup> Foucault contrasts the experience of ‘unreason’ in the “classical age” of madness (i.e. during the late seventeenth and eighteenth centuries), with that of the modern period that followed – from this analysis he argues that the emergence of modern “mental illness” can be explained by a series of social, ethical, political, and philosophical processes that derive from (and, therefore, maintain key elements of) the classical experience of ‘unreason’ and ‘madness’. One of Foucault’s key claims is that “[T]he positivist psychiatry of the nineteenth century, like our own, [...] inherited the relationship that classical culture as a whole had set up with unreason.”<sup>172</sup> To clarify this claim, we need to explain the distinction between ‘unreason’ and ‘madness’ which Foucault deems characteristic of the “classical age”:

According to Foucault, classical “unreason” included not only the mad (i.e. those who will be later called “mentally ill”), but also the poor or unproductive (i.e. the vagabonds, beggars, and the unemployed),<sup>173</sup> and the “immoral” (i.e. venereal sufferers, libertines, homosexuals)<sup>174</sup> who will later constitute the category of the “abnormal”. During the

---

<sup>171</sup> Foucault refers mainly to the changes proposed by Philippe Pinel and Jean-Étienne Esquirol.

<sup>172</sup> Michel Foucault, *History of Madness*, trans. Jonathan Murphy and Jean Khalfa (New York: Routledge, 2006), 159.

<sup>173</sup> *Ibid.*, 81.

<sup>174</sup> *Ibid.*, 83, 88-9.

classical age, Foucault argues, “unreason” was the central worry, not madness. Foucault explains: “...for the classical age unreason had the value of a noun, and had a substantive function. [...] For men of the classical age, madness was not the natural condition, the psychological and human root of unreason, but rather its empirical form.”<sup>175</sup> In contrast, after the birth of modern psychiatry, mental *illness* becomes the prevailing phenomenon and, according to Foucault, it becomes impossible to “confront unreason” without entering the *pathological* dimension of the mad.<sup>176</sup>

Then, when Foucault claims that modern and contemporary psychiatry “inherited” the relation that the classical culture had with the broader category of “unreason”,<sup>177</sup> what is implied is that the modern treatment of the “mentally ill” kept attitudes and practices previously associated with the management of the “unproductive” and the “immoral”. According to Foucault, the latter would explain, for example, an important contradiction between the discourse of modern psychiatry and its practices. Within the positivistic *discourse* of modern psychiatry, Foucault argues, mental alienation was associated with innocence: the mad was depicted as someone subjected to mechanisms that were closer to those of the “natural” world than to those of the moral realm.<sup>178</sup> However, the *treatment* that modern psychiatry gave to the mad abided by different rules, insofar as the mentally ill were subjected to forms of *moral* control.<sup>179</sup> In a word, Foucault claims that modern psychiatry (and “our own”) “thought that madness was purely being studied from the point of view of an objective pathology; but despite those good intentions, the truth was that madness was still haunted

---

<sup>175</sup> Ibid., 156.

<sup>176</sup> Ibid., 352.

<sup>177</sup> Ibid., 159.

<sup>178</sup> Foucault writes: “our age” (i.e. modernity) “has fallen into the habit of conceiving of madness as a fall into a kind of *determinism* where all forms of liberty are slowly eroded” (Ibid., 156; my emphasis). Foucault claims that, while madness was also associated with “animality” in the classical period, it was taken to mean very different things. Madness in the classical age was indeed a form of “animal *freedom*”. (Ibid.; my emphasis) The latter could be read in at least two ways: First, the “animality” of the mad denoted an existence that was *free* from human precariousness and fragility. (Ibid., 148) Second, the mad were also free in the sense that they escaped classical “definitions” of humanity (e.g. “man is as a rational animal”).

<sup>179</sup> Ibid., 159.

by an ethical view of unreason, and the scandal of its animal nature.”<sup>180</sup> Crucially, the latter implies that, the mechanisms put in place to deal with mental illness, are not exclusively ways of treating or “healing” the ill. Additionally, they could be seen as ways of dealing with the socially undesirable effects of unreason and, indeed, as ways of *governing* people.<sup>181</sup>

This Foucauldian thesis can be illustrated through more contemporary examples. For instance, as I mentioned above, one paradigmatic example of a “socially undesirable” effect of “unreason” during the classical age was for Foucault unproductivity – unproductivity threatened economic growth and stability and as a result needed to be dealt with. What could be the effects of conflating ‘mental illness’ with a broader notion of ‘unreason’ in this case? One possible line of thought is considering whether this conflation translates into a tendency to *pathologise* unproductivity. Indeed, some contemporary analyses of mental disorders follow this path:

Consider for example Alain Ehrenberg’s analysis of depression in contemporary societies.<sup>182</sup> Ehrenberg characterises depression as a *style of action* instead of a subjective

---

<sup>180</sup> Ibid.

<sup>181</sup> The reading of the *History of Madness* that I provide takes for granted a continuity in Foucault’s interests that may be seen as disputable. Foucault’s early archaeological writings are typically seen as dealing with epistemological concerns rather than with power. I do not wish to engage in this discussion here. Let me just point out that a radical separation between what can be thought epistemic and its power effects would be very un-Foucauldian. For a description on how to understand epistemic concerns in Foucault see: Beatrice Han, *Foucault’s Critical Project: Between the Transcendental and the Historical* (Stanford: Stanford University Press, 2002), 104; for a discussion on the role of power in the *History of Madness* see Amy Allen, “Feminism, Foucault, and the Critique of Reason: Re-reading the History of Madness.” *Foucault Studies*, no. 16 (September 2013), 15-6. See also Foucault’s own statements: “When I think back now, I ask myself what else it was that I was talking about in *Madness and Civilization*... but power? Yet I am perfectly aware that I scarcely ever used the word and never had such a field of analysis at my disposal.” (Michel Foucault, “Truth and Power,” in *Power Vol. 3 of Essential Works of Michel Foucault 1954-1984*, ed. James D. Faubion (New York: The New York Press, 2000), 117). Furthermore, in the first lecture of his course *Psychiatric Power* he criticises his own reflections in the *History of Madness* and regrets that his analysis focused on violence instead of in apparatus of power. (Foucault, *Psychiatric Power*, 13).

<sup>182</sup> Alain Ehrenberg, *The Weariness of the Self: Diagnosing the History of Depression in the Contemporary Age* (Montreal: McGill-Queen’s University Press, 2010). Admittedly, Ehrenberg’s work could be considered counterproductive for my research. For Ehrenberg “values associated with autonomy” (Ibid., xxx) are part of the social conditions within which depression and, more precisely, “the expansion of the diagnosis of depression” (Ibid., xxviii) become possible. Ehrenberg claims: “I do not approach depression as a weakening of social bonds but, rather, as an attitude, a mindset heavy with multiple social practices and representations of ourselves in a society in which values associated with autonomy (e.g., personal choice, self-ownership, individual initiative) have been generalized.” (Ibid., xxx) I find Ehrenberg’s work compelling and meriting a much deeper analysis that the one I can develop here. However, I do not believe that his analysis applies to *any* conception of autonomy. For example, the feelings of inadequacy, weariness, and void that are central to Ehrenberg’s thesis could be read as



state.<sup>183</sup> One of Ehrenberg's key points is that depression is not merely connected to loss (as the typical view on depression sustains) but is the reaction or symptom of an individual who is "inadequate" with regards to norms of action.<sup>184</sup> Ehrenberg claims: "[t]he ability to act by oneself is at the heart of socialization, and the breakdown of action is the fundamental disorder of depression."<sup>185</sup>

The latter is reflected, according to Ehrenberg, in the psychiatric approach to depression: Ehrenberg notes how antidepressants are meant to "enhance" and "empower" depressed people, which happen to be key positive ideals within a democratic ethos where individuals need to be able to make something of themselves.<sup>186</sup> Ehrenberg claims: "Prozac is not the happiness pill but, rather, the initiative drug".<sup>187</sup> In this view, then, healing or wellbeing are not at the centre of the psychiatric treatment of depression, the *restoration of action* is.

Less provocatively, one could say that a particular conception of healing and wellbeing that fits with a certain social utility (e.g. a conception of wellbeing oriented towards action, entrepreneurship or self-management) is still at the centre of psychiatric practice today. Indeed, the influential *American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorder* (DSM-IV) considers the "impairment in social, occupational, or other important areas of functioning" as appropriate criteria for diagnosing mental disorders.<sup>188</sup> Moreover, "unreasonable" responses to situations, "excessive" forms of doubt and caution, "undue"

---

a (social and personal) "forgetting" of the fact that autonomy has social conditions of possibility and is not something that the individual could or should expect to achieve on her own. What is more, I believe that a model of autonomy that enables to limit oppressive or over-demanding injunctions to self-management could avoid some of the psychic costs that Ehrenberg associates with autonomy. I come back to this point in Chapters 4 and 5.

<sup>183</sup> Ibid., 163-4.

<sup>184</sup> Ibid., xvi.

<sup>185</sup> Ibid., 190.

<sup>186</sup> Ibid.

<sup>187</sup> Ibid., 188.

<sup>188</sup> Fabian Freyenhagen and Tom O'Shea, "Hidden substance: mental disorder as a challenge to normatively neutral accounts of autonomy," *International Journal of Law in Context* 9, Special Issue 1 (March 2013): 57-8. The article refers to DSM IV, but DSM V continues to include similar value-laden characterisations of symptoms.

preoccupation, “inappropriate” behaviour (with regards, for example, to available social norms concerning how to behave given one’s age or culture), among others, may be taken as evidence of different mental disorders.<sup>189</sup> In a word, failure to live up to social ideals sometimes translates into pathologies and, as I argue below, the latter makes tying ‘competence’ to the lack of mental pathologies dangerous in the context of a model of autonomy like the one I aim at.

Let me now move on to an analysis of the ideal of ‘minimal rationality’ more generally. As I have mentioned above, ideals of reflective competence are highly dependent on what a society identifies as *rational* or *reasonable*. Foucault’s analysis is also useful to situate the above-mentioned debates on the “non-neutral” nature of contemporary psychiatric discourses within the broader issue of the separation between “reason” and “unreason” in a given society. Indeed, feminists have long problematized the alleged neutrality of the (broader) notion of “reason” itself. For example, feminists have argued that “reason” has been systematically defined as a “masculine” notion and that, as a result of this, this category may exclude individuals who have not been socialised as (or, in some contexts and historical periods, who have not been given the education or training typically reserved to) ‘men’.<sup>190</sup>

According to Allen, feminists’ critiques of modern reason are divided among those who demand a space for the “other(s)” of reason (i.e.: madness, irrationality, embodiment, affect, femininity, and queerness) within reason and those who claim that reason is something despotic in itself and should therefore be rejected altogether as an ideal (and replaced, for example, by a romantic ideal of unreason or madness).<sup>191</sup> For Allen, the problem with any of these reactions is that they leave unquestioned the division between reason and unreason.<sup>192</sup> According to Allen, a way out of this dilemma would be possible precisely by

---

<sup>189</sup> Ibid., 58.

<sup>190</sup> For a classic example, Genevieve Lloyd, “The Man of Reason,” *Metaphilosophy*, 10, no. 1 (January 1979).

<sup>191</sup> Allen, “Feminism, Foucault, and the Critique of Reason,” 17.

<sup>192</sup> Ibid., 21.

turning to Foucault's considerations on *History of Madness*. Foucault's early writings exhibit a "commitment to engage in a *rational* critique of rationality that highlights reason's dangerous entanglements with power while resisting the temptation to reject or refuse reason altogether".<sup>193</sup>

Allen is against a "simplistic" reading of Foucault according to which "freedom is the embrace of unreason" but defends the idea that "the figure of unreason serves to illuminate lines of fragility and fracture in our historical a priori".<sup>194</sup> That is, what is excluded from the domain of the reasonable (or even deemed pathological) in a sociohistorical context, is useful to signal the limits and contingencies of our historical a priori. In other words, and according to the analysis carried out previously in this chapter, noticing what is labelled as "unreasonable" in a certain sociohistorical context could contribute to the work of critique.

To conclude, the elements presented above warn us against a potentially undesirable effect of the competence conditions embedded in procedural accounts: the transgression of values functional to the social order and, even, the adoption of styles of agency where the "others of reason" (e.g. affective reactions, "disruptive" experiences, embodiment) are given more credit might make an agent fall short of 'reflective competence'. This issue is all the more pressing in light of the aims I have set for my project, insofar as limiting one's heteronomy, in oppressive contexts, might indeed involve challenging socially accepted ways of thinking and being. If opportunities for challenging dominant ways of thinking or lifestyles are restricted from the outset, then a procedural model might not have the critical potential necessary to limit oppressive heteronomy.

I argue that the dangers of indirectly imposing ways of living should be explicitly acknowledged when defining a model of autonomy and I address this issue in the context of

---

<sup>193</sup> Ibid., 18

<sup>194</sup> Ibid., 22.

my critique of historical models in Chapter 3. Moreover, as I argue in Chapter 3, these dangers should be assessed in light of the different interferences that agents might face in different oppressive circumstances. What this means is that competence conditions, in the context of this thesis, should be (loosely) defined having in mind the sets of skills and capacities necessary to *counter* different forms of oppression.

For example, if (as I argue later), limiting oppression might involve questioning long-accepted social “truths” about oneself, then a capacity to imagine oneself otherwise or to put oneself in others’ shoes might become crucial. Conversely, in the latter case, too strong an emphasis on the coherence typically presupposed by conditions of ‘minimal rationality’ and ‘authenticity’ might end up reinforcing oppression. As I have already anticipated, then, the points made so far should not be interpreted as a call to provide models completely free from any normative commitments (such an aim would be, indeed, at odds with defining a model of autonomy) but as an invitation to make sure that competence conditions do not actually curtail agents’ possibilities for limiting oppression.

## II. What Kind of Historical Materials are Necessary for Critical Self-assessment?

In Chapter 1, I have dealt with three different historical models of personal autonomy: Dworkin’s, Mele’s and Christman’s.<sup>195</sup> These three models consider that an agent’s past or, more specifically, the history of a certain content of reflection, is key information when it comes to determining the autonomy of an agent vis-à-vis this content. As I have also shown, the way in which historical information is taken into account varies in these three models.

---

<sup>195</sup> In the existing literature on autonomy, Dworkin’s model is not typically considered a ‘historical’ model. I have argued that it could be considered historical generally speaking insofar as it *does* consider relevant what happened *before* the actual reflective process that determines an agent’s autonomy. In a word, Dworkin’s model is certainly not a structural time-slice model like Frankfurt’s.

To recall: Dworkin and Mele's historical considerations are aimed at determining if an agent's past is free from the kind of interventions that would make autonomy implausible. Christman, on the other hand, makes history a *material* for reflection, i.e. historical information is meant to put agents in a better position to carry out self-assessments. In other words: Christman offers a sophisticated non-paternalist historical condition which makes central the agents' own *interpretation* of their history.

While I deem Christman's move promising for the reasons explained in Chapter 1, in this section I want to start to make the case that, if a historical model is to serve the aims I set for my project, then two key questions need to be further problematized – namely: (i) what *kind* of history should inform self-assessment? and (ii) what kind of *reading* of one's history should inform self-assessment?

Regarding (i), as I have explained in Chapter 1, the *scope* of Christman's historical condition is not sufficiently clear. Throughout Christman's work, the historical condition in his model appears to be sometimes understood narrowly (e.g. as one's "psychological history" alone) and other times more broadly (e.g. as "social" history). Moreover, even granting that Christman's model is indeed flexible enough to accommodate a broader kind of historical condition (which surely seems in line with Christman's focus on *sociobistorical* selves), it is still necessary to provide an analysis of what *justifies* understanding history in one way or the other in the context of this thesis.

In what follows, I turn to the notion of 'genealogy'. I analyse the notion by focusing mostly on Foucault's reading of Nietzsche but I also turn to the work of contemporary theorists who reflect on 'genealogy' in a Foucauldian way. I proceed as follows: First, I analyse the particular type of history-telling that genealogy implies. Second, I analyse the practical and even "existential" effects that genealogy has on those whose histories are "revised" through genealogical critique. Third, I show that genealogies cannot be reduced to "factual" histories (i.e. to histories as a succession of events) but that, rather, they need to be

understood as histories of sociohistorical conditions of possibility or of social determinations.

It should be noted that I present below what I consider are key features of genealogy given my research on personal autonomy, but I do not intend to be exhaustive in my presentation. The resources drawn from the discussion carried out in this section will help me to refine my two worries vis-à-vis Christman's historical conditions and to carry out a more developed critique of Christman's model in Chapter 3.

Let's begin by framing 'genealogy' vis-à-vis 'history': indeed, as it will become clearer throughout this section, the specificities of genealogical reflection typically emerge from *contrasting* genealogy and (other kinds of) history. That is: both genealogy and history study an object's "past" (and, when this is applicable, the way in which this past affected or ended up configuring the object in the present), but they do so differently.

A first approximation to genealogy could be put as follows: genealogy is a "non-directed" form of historicism. Foucault argues that a genealogist should avoid *imposing* a meaning on history through meta-historical ideals like teleology (that is: believing that history has a clear and indisputable direction or *telos*) or through the myth of a "true" origin of the present to be found in the past.<sup>196</sup> It is not the task of the genealogist to look for an original sense which could have been lost (or made less-visible) as history went by. Neither should a genealogist attempt to re-establish the "true" meaning of history. Rather, for Foucault, genealogies aim at constructing "effective histories" which do not *presuppose* big continuities, stabilities, or a *telos*.<sup>197</sup> For example, genealogies are typically contrasted to "commonplace" developmental histories of the nineteenth century which would see societies and human life evolving or progressing according to specific principles – Comte, Marx, and Hegel are

---

<sup>196</sup> Foucault, "Nietzsche, la Généalogie, l'Histoire," in *Hommage à Jean Hyppolite*, ed. Suzanne Bacherlard et al. (Paris: Presses Universitaires de France, 1971), 152.

<sup>197</sup> *Ibid.*, 160.

examples of the latter.<sup>198</sup> Insofar as genealogies avoid these forms of *orienting* narratives or *meta-histories* genealogies would constitute a form of “radical historicism”.<sup>199</sup>

Furthermore, the genealogist’s move is also to question that there is such a thing as an objective and incontestable origin of objects or practices. Foucault goes as far as to argue that interpretation (of history, of meanings) “is a never-ending task [...] because there is nothing to interpret”.<sup>200</sup> Instead, one works already on other interpretations. Without the need to commit to a general relativist view (which I do not intend to do), the point on history being a mainly *interpretative* task is one that merits consideration:

For one thing, the point about history being always an interpretation could be read as a “methodological warning”, so to speak. A genealogist should be aware of the interpretative nature of the narrative that she constructs insofar as historical narratives typically involve, for example, selecting facts or imposing coherence and continuity where there is no necessarily such thing. So, anyone carrying out a historical inquiry should be wary of the fact that reading a history will imply adopting a perspective among others and that, thus, deciding or making explicit “where” one is judging from is at least as important as identifying a historical object for analysis.<sup>201</sup>

However, the point about history being an interpretative task is not merely a methodological one or an epistemological one.<sup>202</sup> When a genealogist considers different possible readings of the same history, she should also be aware of the fact that favouring a particular reading is *not* a neutral decision to make. Take Nietzsche’s genealogy of morality as an example. The latter can be read as more than just an attempt to show that there could

---

<sup>198</sup> Mark Bevir, “What is Genealogy?,” *Journal of the Philosophy of History* 2 (2008): 266

<sup>199</sup> Ibid.

<sup>200</sup> Foucault in Hubert L. Dreyfus and Paul Rabinow, *Michel Foucault: Beyond Structuralism and Hermeneutics*. (Chicago: The University of Chicago Press, 1983), 107.

<sup>201</sup> Foucault, “Nietzsche, la Généalogie, l’Histoire,” 163.

<sup>202</sup> According to Amia Srinivasan, genealogies do not merely (or, even, mainly) question the “epistemic standing” of our representations but, rather, they focus on “what our representations do – and in what we might do with them.” (Amia Srinivasan, “Genealogy, Epistemology and Worldmaking,” *Proceedings of the Aristotelian Society* 119, Part 2 (2019): 141).

be a different history of our morality. It also intends to expose how *tainted, low, or interested* some of these histories are.<sup>203</sup>

This takes me to my second point: carrying out a genealogical critique of one's values has effects on our experience vis-à-vis those values. That is: genealogists suggest that by learning about the historical, contingent, and power-affected origin of our values, we may come to see these values differently. For example, Bevir claims: "when other people believe that certain social norms or ways of life are natural or inevitable, radical historicists [i.e. genealogists] denaturalize these norms and ways of life by suggesting that they arose out of contingent historical contexts."<sup>204</sup> According to Bevir, genealogies enhance one's freedom because, by understanding that one's current way of being is not necessary, the possibility of creating and implementing new practices opens up. Indeed, Bevir adds: "..., genealogy opens novel spaces for personal and social transformation precisely because it loosens the hold on us of entrenched ideas and institutions; it frees us to imagine other possibilities"<sup>205</sup>.

It is not hard to see how "opening up novel spaces for personal and social transformation" could already be of relevance for a model of autonomy like the one I aim at. However, I want to push this point a bit further here. To put it differently, my worry now is the following: is genealogical analysis "just" another way of guaranteeing the "modal change" (i.e. seeing previously-thought "necessities" as "contingencies") that we analysed when discussing the 'critical ethos'? I want to argue that this is not the key contribution of genealogical reflection and this for at least three reasons:

First, genealogical critique does not stop at showing the contingency of practices. Additionally, genealogical critique has a "loosening" effect thanks to which the practices we identified so clearly reappear as more incoherent or disjointed. The fact that these practices

---

<sup>203</sup> Friedrich Nietzsche, *On the Genealogy of Morals*, trans. W. Kaufmann (New York: Vintage Books Editions, 1989), 20.

<sup>204</sup> Bevir, "What is Genealogy?," 271.

<sup>205</sup> *Ibid.*, 272.



could be “opened” for us to intervene on them is not just a point about the contingency of the practices themselves, but also about the arbitrariness of our taxonomy of different practices.<sup>206</sup>

For example, a genealogical revision of some of our ethical practices could shed light on the fact that monitoring one’s desires from close, and always keeping their intensity at bay so that they do not “revolt” against reason, might not be the only or the necessary approach to morality. One could understand the contingency of the model and decide, so to speak, to give desires more credit. However, opening this practice of morality to change, also means questioning why self-surveillance, desires, and reason appear to be the natural elements of what we call a moral practice. A genealogical approach would also show that in the very gesture of designing a cluster of practices as the necessary ensemble which constitutes, say, moral experience or human sexuality, there is already the imposition of a coherent unity over discrete phenomena. So, the critical power of genealogy here lies not only in showing that our objects of study (ourselves, institutions, etc.) could have been or could be otherwise, but also in identifying these objects differently to begin with – notably as less coherent than they appear.

Second, genealogical analysis provides more than an unveiling of “contingencies” because of my earlier point on the “tainted” origin of some of our historical interpretations. Foucault’s reading of Nietzsche, for example, suggests that interpreting history always implies an “interested” appropriation of systems of rules and meanings to impose a direction on them.<sup>207</sup> Genealogy is meant to tell the story of these appropriations and, in this sense, it could be considered a critical device to unveil the *power-affected* ways in which historical objects are (self-)constituted.

---

<sup>206</sup> See Raymond Geuss, “Genealogy as Critique,” *European Journal of Philosophy* 10, no.2 (December 2002) on this point. Geuss argues that genealogies, particularly in Foucault, “offer a historical dissolution of self-evident identities”. (Ibid., 212).

<sup>207</sup> Foucault, “Nietzsche, la Généalogie, l’Histoire,” 158.

Therefore, by insisting on contingency too much, we might forget that our practices, identities, or institutions, did not just end up being the way they are by “flipping a coin”, so to speak. Problematizing an object through genealogical critique also means understanding *why* a particular configuration prevailed, to serve what purposes, to enable what other practices, identities or institutions. In other words: it is not only the lack of metaphysical necessity that is at stake, but also a certain social “purposiveness” which makes historical objects what they need to be given cultural, social, economic or political aims. As Amia Srinivasan puts it, the crucial question for “critical genealogists” (among whom she includes Foucault, Simone de Beauvoir, bell hooks, Angela Davis, Catharine MacKinnon and Judith Butler) is “[w]hat practices and forms of life do [our representations] help sustain, what sort of person do they help construct, and whose power do they help entrench?”<sup>208</sup>

Third, genealogies do more than unveiling contingencies if by the latter we picture a purely “intellectual” grasp of this contingency: “effective” histories have, according to Foucault, *disruptive* effects on our emotions, instincts, and bodies.<sup>209</sup> This gives new sense to my earlier claim on genealogies being able to modify our “experience” vis-à-vis the historical objects under analysis. According to Martin Saar, for example, genealogical texts are different from “mere methodological historicism”<sup>210</sup> insofar as the former are “‘practical’ texts with an existential dimension.”<sup>211</sup> Saar explains:

The aim of genealogical writing is exactly to raise affects and to stir up doubts and questions about the present form of subjectivity. One might say that the genealogical text is meant to put the readers’

---

<sup>208</sup> Srinivasan, “Genealogy, Epistemology and Worldmaking,” 142.

<sup>209</sup> Ibid., 160; 168.

<sup>210</sup> Martin Saar, “Understanding Genealogy: History, Power, and the Self,” *Journal of the Philosophy of History* 2 (2008): 298.

<sup>211</sup> Ibid., 311. Saar’s considerations refer to a genre of “texts,” that is: to a particular way of writing history. I use Saar’s considerations to illustrate something about a more general genealogical ‘style’ which might be useful to discuss genealogical ways of, not just writing, but also telling, considering, and reflecting on history.

identity into crisis by confronting them with descriptions about themselves that radically contradict their own self-understanding and thereby to encourage them to revise their judgements and practices, and this means ultimately, to revise themselves.<sup>212</sup>

For Saar, the uneasiness or discomfort which genealogy generates in those to whom the narrative is addressed and whose ways of being are problematized, is exactly the *aim* of genealogy. The fact that historical subjects become aware of the power-affected and socially-mediated ways in which they constitute and/or understand themselves facilitates rejection and estrangement and, eventually, self-revision.

I conclude this section by introducing the third feature of genealogy that I anticipated above, namely how genealogy deals with histories that are more fundamental than biographical histories or, in other words, that are more *social* than *personal*.

For example, in Saar's text above, it is quite clear that the "identity" at stake should not be understood as a synonym of "character" or "personality". Indeed, when authors connect genealogy with the possibility of questioning "ourselves", what is being questioned is typically a more "collective" identity, a sociohistorical way of being, i.e. a 'subjectivity' (in Saar's words).

Then, the question arises of which historical objects are suitable candidates for a genealogical analysis which brings about "existential" effects. What I mean is that, to achieve the latter aim, genealogies should not analyse any historical object but rather those objects somehow relevant to individual identities. Indeed, when Foucault identifies possible "domains of genealogy", he identifies three historical 'axes' which allow to study the way in which we become different kinds of *subjects* – "subjects of knowledge", "subjects acting on others" (i.e. subjects within power relations), and "moral agents".<sup>213</sup>

---

<sup>212</sup> Ibid.

<sup>213</sup> Foucault, "On the Genealogy of Ethics," 262.

To put it less abstractly, genealogies are – as Saar suggests – critiques of those practices, institutions, and concepts which direct individuals’ conducts by acting on the available ways through which these individuals can understand themselves and relate to themselves. Saar claims:

Genealogies are stories told about the historical emergence and transformation of concepts, practices or institutions that relate to the making of selves by influencing their self-understanding and way of conduct. [...], these stories narrate histories of the self as histories of power.<sup>214</sup>

It seems then that models of the person, “kinds of people”, socially available character traits, could be considered apt candidates for a genealogical inquiry. These categories affect the ways in which agents understand and define themselves and, most importantly, they direct agents’ conducts by channelling their efforts to fit into socially available categories.

To summarize, the reflections on genealogy that I have presented in this section allow me to develop the two different (though related) questions presented at the beginning of this section. (To recall: (i) what *kind* of history should inform self-assessment? and (ii) what kind of *reading* of one’s history should inform self-assessment?).

Regarding (i), I can begin to make the case that historical models of personal autonomy should include a historical condition which is explicitly *social* if they are to be sensitive to power-affected ways of being which increase agents’ heteronomy. In other words, if we want agents to have opportunities for questioning, for example, social injunctions which define socially available characters, then the materials provided by personal histories alone will not take us far enough. What reappears here is then the need to consider the history of one’s

---

<sup>214</sup> Saar, “Understanding Genealogy,” 307.

contemporaries, of those subjected to the same kind of historical determinations. The understanding one could gain from considering one's biography in isolation would not be sufficient to achieve the aims that I have presented in this section.

Regarding (ii), the difference between genealogy and history helps us understand and define the risks of taking history or historical materials as "givens". In particular, historical models of autonomy should consider the possibility that certain "directed" readings of one's history (e.g. those directed by ideals of authenticity) could end up reinforcing certain oppressive circumstances. To be sure, it would be over demanding to expect that all agents become critical genealogists to be autonomous and this is not my point here. My suggestion is that the resources analysed above should be taken into account by autonomy *theorists* when defining *models* of autonomy capable of dealing with oppression.

Finally, it is worth noting that genealogical critique might be promising to revise historical models insofar as it problematizes the reflective and affective conditions necessary for experiencing certain forms of "existential" discomfort. In Chapters 3 and 5, I connect this point with my reflection on (what contemporary theorists of autonomy call) 'alienation'. One of my points is that, if the absence of experiences of alienation is to be taken as a meaningful criterion, this needs to happen under conditions in which feeling alienated is actually a possibility for the subjects involved. As I argue in Chapter 3, I do not believe that models where self-assessment is carried out in light of the 'authentic' status of one's sustained values or settled traits of character provide agents with sufficient opportunities to experience 'alienation'.

I unpack and justify these three points in the chapters that follow: when developing my critique of historical accounts (in Chapter 3) and through the examples and study-cases that I present in Chapters 4 and 5.

### III. How Should We Understand Oppression?

As I mentioned in the introduction to this chapter, procedural-historical models seem to assume that potentially illegitimate heteronomous interventions happen according to a disruptive schema, i.e. that one can typically identify a moment in which this kind of intervention *interrupted* an agent's "normal" development. However, this assumption might leave unproblematized sustained developments which may be themselves shaped by oppression. Moreover, I have argued that taking deception and manipulation as paradigmatic interferences assumes that there is always a clear "other" who dominates or deceives the agent whose autonomy or heteronomy is under analysis. The latter also seems inadequate to capture common forms of oppression, as I have discussed above.

What is still missing at this point is a discussion on the ways in which oppression typically operates. I turn once more to Foucault to provide an alternative schema to analyse different forms of oppression:

As I mentioned in the Introduction, though Foucault rarely uses the term "oppression", his analysis can accommodate not only general cases of illegitimate heteronomous interferences like the ones mostly discussed by proceduralists (e.g. some forms of manipulation or violent interventions) but also others that substantive theorists of autonomy would consider oppressive (e.g. being problematically shaped by an existing moral code).

My aim in this section is to distinguish between two different phenomena: 'domination' and 'subjection'. In order to show both how this distinction operates in Foucault's work and to explain my own interpretation of this distinction, I need to discuss some passages from Foucault's late interviews closely. This kind of exegetical work is in this case justified by the key role which this Foucauldian-inspired distinction plays in the two-tracked model of autonomy which I develop in Chapter 5.

Foucault defines power relations as “strategies by which individuals try to direct and control the conduct of others” or as “strategic games between liberties”.<sup>215</sup> The first thing to note is then that, as I mentioned in the Introduction, ‘power’ is not a thing or a substance: Foucault analyses “power *relations*”, namely relations in which participants try to influence each other’s behaviour in light of a desired outcome.<sup>216</sup> Second, we should note that power involves ‘conducts’ and ‘liberties’. In what follows, I explain these terms by analysing power’s constitutive elements and, crucially, by distinguishing power from ‘relationships of violence’ and from ‘states of domination’.

In the majority of his work, Foucault does not clearly distinguish between “power” and other notions like “violence”, “coercion”, and “domination” – indeed Foucault frequently uses these terms interchangeably.<sup>217</sup> However, during the 1980s Foucault tries to capture more carefully what constitutes the *specificity* of power relations vis-à-vis these (related) notions.

Foucault claims that two elements are “indispensable” for us to be able to identify a power relation:

..., a power relationship can only be articulated on the basis of two elements which are each indispensable if it is really to be a power relationship: that “the other” (the one over whom power is exercised) be thoroughly recognized and maintained to the very end as a person who acts; and that, faced with a relationship of power,

---

<sup>215</sup> Foucault, “The Ethics of the Concern of the Self as a Practice of Freedom,” 298; 299.

<sup>216</sup> *Ibid.*, 291-2.

<sup>217</sup> This is notably the case in Foucault’s lectures at Dartmouth College (Michel Foucault, *About the beginning of the hermeneutics of the self: Lectures at Dartmouth College, 1980* (Chicago: The University of Chicago Press, 2015).

a whole field of responses, reactions, results, and possible inventions may open up.<sup>218</sup>

Power relations, then, act on “conducts” insofar as they are exercised over *persons* (i.e. not on objects) who can act (i.e. conduct *themselves*) in significantly different ways (i.e. who have a “whole field of responses” available).

In other words, Foucault distinguishes the *indirect* “conduct of conducts” which is (in Foucauldian terms) a distinctive marker of power, from the more *direct* interventions on people’s bodies and actions (e.g. through the use of coercion or sheer violence) that people ordinarily associate with power.<sup>219</sup> Power is not about “breaking” bodies but, rather, about “inciting”, “inducing”, and “seducing” conducts.<sup>220</sup> We can say, in short, that power relations *influence* the way in which *agents* exercise their *freedom*.<sup>221</sup>

“Freedom”, then, is a condition of possibility of power: when individuals participate in strategic power relations, they are not *forced* (or they do not force others) to act or conduct themselves in one way or another. In addition to being forced by bodily or physical constraints (e.g. by the exercise of sheer violence, as explained above), one can be forced to act in a certain way if a particular state of affairs is made the *only* available option. The exercise of power is a “management of possibilities”<sup>222</sup> which implies that relationships of power are senseless when “determining factors are exhaustive” as it happens, for example, when only one course of action or state of affairs is possible.<sup>223</sup> Even more, that power is exercised on individuals who are *free* to some extent, seems to mean in this context that individuals playing

---

<sup>218</sup> Foucault, “The Subject and Power,” 340.

<sup>219</sup> Foucault, “The Ethics of the Concern of the Self as a Practice of Freedom,” 299. I do not imply that power does not act on bodies at all, ‘discipline’ (as explained above) does indeed. However, there is a difference between, for example, monitoring and examining bodies with the purpose of normalizing them, and violently breaking them.

<sup>220</sup> Foucault, “The Subject and Power,” 340-1.

<sup>221</sup> *Ibid.*, 340.

<sup>222</sup> *Ibid.*, 341.

<sup>223</sup> *Ibid.*, 342.



the game of power are capable of *meaningfully* different possible actions. Let me unpack and justify this claim:

Foucault suggests that the mere existence of some however minimal freedom (i.e. the existence of *any* space of possibility for action and in *any* extent) is not enough for power to be possible. Consider Foucault's claim on slavery, for example: "slavery is not a power relationship when a man is in chains, only when he has some possible mobility, even a chance of escape."<sup>224</sup> Mentioning the chance of escape here is, I believe, highly relevant: the freedom of mobility necessary within power relations should allow to *reverse* the power relation. Foucault makes this point clearer when he claims:

... at the heart of power relations and as a permanent condition of their existence there is an insubordination and a certain essential obstinacy on the part of the principles of freedom, then there is no relationship of power without the means of *escape* or possible *flight*. Every power relationship implies, at least in potentia, a strategy of struggle, in which the two forces are not super-imposed, do not lose their specific nature, or do not finally become confused. Each constitutes for the other a kind of permanent limit, a point of possible reversal.<sup>225</sup>

Therefore, when one of the parties within a power relation is reduced to an object (i.e. annihilated as a subject) or when she is forced to act in a specific way (e.g. because the margin of freedom available is extremely limited and, crucially, it does not allow for the reversal of the power relation from within), we are no longer confronted with a power relation. When,

---

<sup>224</sup> Ibid., 342. A similar point is made in at least another occasion: "The characteristic feature of power is that some men can more or less entirely determine other men's conduct - but never exhaustively or coercively. A man who is chained up and beaten is subject to force being exerted over him. Not power. But if he can be induced to speak, when his ultimate recourse could have been to hold his tongue, preferring death, then he has been caused to behave in a certain way. His freedom has been subjected to power. He has been submitted to government." (Foucault., "Omnes et singulatim," 324)

<sup>225</sup> Foucault, "The Subject and Power," 346; emphases added.

instead of “inducing or seducing” conducts, bodies are “broken” or spaces of possibility for action are “closed off”, then we face a “relationship of violence”, not a power relation.<sup>226</sup>

Foucault uses the same criterion of “reversibility” (i.e. that, for power to be such, it needs to be reversible through the margin of freedom available to the subjects within the power relation) when distinguishing power relations from ‘states of domination’. In this case though, Foucault appears to be working on a different level: he does not refer to instances of (so to speak) “micro-violence” like the one’s described above (e.g. inflicting pain on an individual) but rather to forms of “blocking” power relations which affect social groups, casts, or economic classes.<sup>227</sup> Foucault writes:

... one sometimes encounters what may be called situations or states of domination in which the power relations, instead of being mobile, allowing the various participants to adopt strategies modifying them, remain blocked, frozen. When an individual or social group succeeds in blocking –by means which can be economic, political, or military– a field of power relations, immobilizing them and preventing any reversibility of movement, one is faced with what may be called a state of domination.<sup>228</sup>

Foucault gives the example of the (European) “conventional marital structure of the eighteenth and nineteenth centuries” as another case in which the “few options” that women had were ultimately only “stratagems that never succeeded in reversing the situation”.<sup>229</sup> The key point is that, under states of domination, “practices of freedom do not exist or exist only unilaterally or are extremely constrained and limited” and, therefore, certain operations of

---

<sup>226</sup> Ibid., 340.

<sup>227</sup> Ibid., 348.

<sup>228</sup> Foucault, “The Ethics of the Concern of the Self as a Practice of Freedom,” 283; translation modified.

<sup>229</sup> Ibid., 292.

(economic, legal or political) *liberation* (e.g. from “male power”) are needed *before* power can be countered through more individual strategies.<sup>230</sup>

‘States of domination’ are different from the ‘relationships of violence’ described above for at least two reasons:

First, because (as I have already mentioned), ‘states of domination’ concern ossifications of structural or systemic power relations and not exclusively of those power relations which could happen on a micro level. To be sure, the micro and the macro levels could be connected and Foucault indeed seems to suggest that they can be: Foucault’s example of the violence exerted over a slave could be considered a case of a relationship of violence which derives from a situation of systemic domination (e.g. within the context of chattel slavery).

Second, ‘states of domination’ are not necessarily reducible to ‘relationships of violence’ because, in examples like the one concerning the European marital structures of the eighteenth and nineteenth centuries, we could not say that women’s space of possible action were “closed off” as they would be in a ‘relationship of violence’. Wives in this context were, indeed, agents acting in different forms and developing different strategies to navigate their situations. What is crucial, however, is that the margin of freedom available to them, their room for play, was not such that their situation could be reversed “from within” their state of domination. As McWhorter puts it, in this case proposed by Foucault:

Male domination in marriage was neither complete nor permanent, and some feminine resistance to particular acts of masculine assertion were likely commonplace [...]; even so, domination was very real, for that regime of power/knowledge could perpetuate itself indefinitely regardless of what individual women might do.

---

<sup>230</sup> Ibid., 283.

Change, when it finally came, did not come—and could not have come—entirely as a result of action from within the dominated group.<sup>231</sup>

Understanding ‘domination’ as a power configuration which cannot be reversed “from within” allows us to capture other phenomena which happen on the *micro*-level and which Foucault also calls ‘domination’. For example, Foucault considers a pedagogical relation between a teacher and a student. According to Foucault, power relations are acceptable in a pedagogical context (e.g. teachers “direct” students in many ways and this is not necessarily a bad thing) but, however, we should avoid “the kind of domination effects where a kid is subjected to the arbitrary and unnecessary authority of a teacher” or where “[the student is] put under the thumb of a professor who abuses his authority.”<sup>232</sup> In these cases, however, the “domination” which could arise is neither one that necessarily reflects macro-level domination nor is exactly a synonym of sheer violence. Rather, Foucault’s use of ‘domination’ here refers to a case where power cannot be played “freely” because of a radical asymmetry of power (i.e. the student is, in a sense, “powerless” or is at least way less powerful than their teacher) which happens on a micro-level.

From now on, my use of ‘domination’ will follow the broader usage described just above. That is, I use ‘domination’ to refer to irreversible, i.e. “frozen”, power relations which could happen both on the micro and on the macro level. Indeed, interpreting ‘domination’ in this sense seems compatible with (and helps to make better sense of) Foucault’s “maxim” that we should play the games of power “with as little domination as possible.”<sup>233</sup> As I see it, this principle could then apply both to the “micro” level of strategic relations between

---

<sup>231</sup> McWhorter, “Post-Liberation Feminism and Practices of Freedom,” 58.

<sup>232</sup> Foucault, “The Ethics of the Concern of the Self as a Practice of Freedom,” 299.

<sup>233</sup> *Ibid.*, 298.

individuals and to the “macro” level of government of populations (presented earlier in this chapter and further defined below). I say more about this in Chapter 5.

To summarize then, ‘relationships of violence’, ‘states of domination’, and ‘domination’ more generally understood (as defined above) can be distinguished from strategic power relations on a crucial point: they are (or have become) ossified and/or they involve the use of direct coercion (e.g. violence) and, as a result, there is little or no room for “playing” the games of power. This distinction is crucial for my analysis of the possibility of limiting oppression through procedural means in Chapter 5. For now, let me anticipate that, while procedural models as they stand are generally sensible to relationships of violence and perhaps to some past interferences tending to “domination” (in the micro-level – i.e. non-systemic – sense suggested above), they need revision to better deal with more systemic forms of domination (i.e. with ‘states of domination’ proper) and with other forms of institutionalised power relations, i.e. with ‘government techniques’. I discuss this latter kind of power relation in the remainder of this section.

In the same interview where ‘states of domination’ are defined, Foucault makes a three-level distinction between (1) dynamic strategic relations (i.e. open-ended power relations), (2) techniques (or technologies) of government, and (3) states of domination.<sup>234</sup> These three levels could be distinguished, as I mentioned earlier, along (at least) two dimensions: i) the amount of individual freedom that they presuppose (it is abundant in strategic relations, present in techniques of government, and almost non-existent in states of domination); and ii) the extent to which they are reversible or dynamic (this includes notably the extent to which they can be altered or reversed by using the margin of freedom available to the

---

<sup>234</sup> Ibid., 299.

individuals within these relations). I have already analysed strategic relations and states of domination, let me say a bit more about techniques of government:

To add to my more general presentation of ‘government’ (in section I. a.), in this section I want to connect the notion of ‘government’ with another aspect of the “problem of individuality”, as I put it earlier. Foucault claims:

Governing people, in the broad meaning of the word, is not a way to force people to do what the governor wants; it is always a versatile equilibrium, with complementarity and conflicts between techniques which assure coercion and processes through which the self is constructed or modified by himself.<sup>235</sup>

In light of the distinctions presented earlier, “coercion” here should be interpreted with caution. In this occasion, when Foucault speaks of “coercion-technologies” or “technologies/techniques of domination”, he is using “coercion” and “domination” as synonyms of “power”. That is, Foucault refers in this case to social practices aimed at directing individuals’ conducts.

Government, then, is a combination of “techniques which permit one to determine the conduct of individuals, to impose certain wills on them, and to submit them to certain ends or objectives”<sup>236</sup> and “technologies of the self”, described as follows:

...techniques which permit individuals to effect, by their own means, a certain number of operations on their own bodies, on their own souls, on their own thoughts, on their own conduct, and this in a manner so as to transform themselves, modify themselves, or to

---

<sup>235</sup> Foucault, *About the beginning of the hermeneutics of the self*, 26.

<sup>236</sup> *Ibid.*, 24-5.

attain a certain state of perfection, of happiness, of purity, of supernatural power, and so on.<sup>237</sup>

Foucault's key point is, therefore, that government cannot be reduced neither to sheer violence (e.g. governing is a way to exert *power* after all) nor to an exclusively "external" form of power, i.e. government also involves directing *self*-relations.

Crucially, as I anticipated earlier, when we say that (modern) governmental techniques are "individualizing" (as well as "totalizing"), this needs to be understood not merely in the sense that each (numerical) individual is concerned, but also in the sense that different "individualities" are concerned (this time in the sense of different "identities", "subjectivities"). Indeed, Foucault defines the 'government of individualisation' (which evolved from pastoral power and characterises Western contemporary societies) as follows:

This form of power that applies itself to immediate everyday life categorizes the individual, marks him by his own individuality, attaches him to his own identity, imposes a law of truth on him that he must recognize and others have to recognize in him.<sup>238</sup>

As Foucault famously claimed, being a "subject" of a certain sort implies a double "tie": one is both tied to others and to *oneself* (e.g. to one's "conscience", to an identity, to certain accepted "truths" about oneself). Moreover, these two modes of "subjection" are intimately related in modernity insofar as, as I have explained earlier, within the modern state the integration of individuals is conditioned on the fact that their individuality is "submitted to a set of very specific patterns".<sup>239</sup> In other words, the 'government of individualisation' which appears after modernity is a form of power that "ties the individual to himself and *submits him to others in this way*".<sup>240</sup>

---

<sup>237</sup> Ibid., 25.

<sup>238</sup> Foucault, "The Subject and Power," 331.

<sup>239</sup> Ibid., 334.

<sup>240</sup> Ibid., 331.

However, if – as I explained earlier – government is a form of power, then the forms of “submission” and “subjection” to which we are referring to should not be presented as irreversible or as assuming an entirely passive picture of the subject within the process of self-formation. Indeed, according to Foucault, the realm of the (directed) self-relations involved in self-constitution and in government – i.e. the ‘ethical’ realm – always presupposes a degree of *freedom*.<sup>241</sup> Let me justify this claim briefly:

Foucault defines a ‘moral code’ as a “set of values and rules of action that are recommended to individuals through the intermediary of various prescriptive agencies [*appareils prescriptifs*] such as the family (in one of its roles), educational institutions, churches, and so forth”.<sup>242</sup> This set of values and rules “legislate” not merely agents’ conducts but also the “relationship you have to yourself when you act.”<sup>243</sup> Actions motivated by the moral code that *fall back on the self*, concern (what Foucault calls) the ‘ethical’ realm. ‘Ethics’, Foucault explains, concerns “the manner in which one ought to form oneself as an ethical subject acting in reference to the prescriptive elements that make up [a moral] code.”<sup>244</sup> In other words, ethics prescribes different ‘modes of subjectivation’.

Crucially, ‘ethics’ (as this particular form of self-relation) is for Foucault a ‘practice of freedom’: it is both made possible thanks to freedom and is the form that freedom takes. Foucault claims: “Freedom is the ontological condition of ethics. But ethics is the considered

---

<sup>241</sup> I follow Arnold Davidson in reading ‘ethics’ as a key element of ‘governmentality’. Davidson explains that ‘ethics’ is “part of both the history of subjectivity and the history of governmentality. Our “technologies of the self,” the ways in which we relate ourselves to ourselves, contribute to the forms in which our subjectivity is constituted and experienced, as well as to the forms in which we govern our thought and conduct. We relate to ourselves as specific kinds of subjects who govern themselves in particular ways. In response to the questions “What kinds of subjects should we be?” and “How should we govern ourselves?,” Foucault offered his history of ethics.” (Arnold Davidson, “Ethics as Ascetics: Foucault, the History of Ethics, and Ancient Thought,” in *Cambridge Companion to Foucault*, ed. Gary Gutting (Cambridge: University Press, 2005), 127)

<sup>242</sup> Michel Foucault, *The Use of Pleasure*: Volume 2 of *The History of Sexuality* (New York: Vintage Books, 1990), 25.

<sup>243</sup> Michel Foucault, “The minimalist Self,” in Michel Foucault, *Interviews and other writings 1977-1984* (London: Routledge, 1988), 15.

<sup>244</sup> Foucault, *The Use of Pleasure*, 26.



[*réfléchie*] form that freedom takes when it is informed by reflection.”<sup>245</sup> Indeed, it is the freedom that is always involved in self-transformation that makes it something which can be governed and that can be integrated with power-technologies for that purpose. This distinction between ethics and morality plays a key role in my critique of substantive accounts, particularly of Natalie Stoljar’s feminist intuition (Chapter 4).

I argue that Foucault’s analysis of forms of ‘government of individualisation’ is useful to analyse those forms of oppression which affect agents’ developments and lives as a whole. For example, even when Foucault did not deal in detail with the production of feminine bodies and subjects in particular, it is possible to think of some aspects of contemporary gender oppression along these Foucauldian lines.<sup>246</sup> As Sandra Lee Bartky points out: “[w]omen have their own experience of the modernization of power” but, still, this experience “follows in many respects the course outlined by Foucault”.<sup>247</sup> Like the processes that Foucault did describe, the processes that play a role in the production of feminine individuals are “dispersed and anonymous”, “invested in everyone and in no one in particular”, and they do not rely “upon violent or public sanctions”.<sup>248</sup>

Then, if we concede (as many feminists do) that certain gender norms are oppressive (e.g. those involving injunctions to subservience or those that link women’s self-worth to motherhood or to certain aesthetic ideals), we have a clear (and common) example of a form of oppression with the following characteristics: (i) it affects an agent’s “normal” and sustained development, i.e. it is a *non-disruptive* form of oppression; (ii) it is dependent on

---

<sup>245</sup> Foucault, “The Ethics of the Concern of the Self as a Practice of Freedom,” 284.

<sup>246</sup> I say “*some* aspects of *contemporary* gender oppression” because, as it emerges from my analysis above, in other sociohistorical configurations, gender oppression may qualify as a ‘state of domination’ (i.e. women are not oppressed now in the same way as women within the European 18<sup>th</sup>-century marital structure were). Moreover, as it becomes clear in Chapter 5, I recognise that one may be, in a specific context, generally ‘subjected’ as a woman but face situations of ‘domination’ if, e.g., should one face an unwanted pregnancy, one’s margin of action would be severely limited by a law that penalises abortion.

<sup>247</sup> Sandra L. Bartky, “Foucault, Femininity, and the Modernization of Patriarchal Power,” in Bartky, *Femininity and Domination*, 79; emphasis added.

<sup>248</sup> *Ibid.*, 79-80.

external injunctions to submit to socially available identity categories, social roles, and values that are (also) *self-imposed*; (iii) it has a force on agents and this without the need of being “violently” imposed on them – rather, they are compelling because they are perceived as “truths” on who we are or how we should live. Moreover, (iv) there is not necessarily a clearly identifiable deceiver or manipulator. Everyone — men and women — are complicit in upholding the gender norms, and perhaps no one upholds them with the intention to dominate, manipulate, or deceive.

Admittedly, likening some forms of oppression with the phenomenon of ‘government of individualisation’ makes the assessment of agents’ autonomy a very complex matter. For example, an agent could be heteronomous – i.e. subjected to others – *precisely* in virtue of her being self-managed. Still, if submission to others is enabled by certain forms of self-management that are *reversible*, then challenging specific forms of self-management can make agents less subjected to others, which means: less heteronomous. I analyse this possibility in Chapters 4 and 5.

Understanding oppression along these lines has, however, challenging consequences for procedural-historical autonomy theories. As I argue in Chapter 3, in spite of their relevance in contemporary liberal societies, historical models are not well-equipped to deal with non-disruptive forms of oppression likened to forms of ‘government of individualisation’. The Foucauldian resources introduced in this section can help to identify the elements in historical procedural accounts which need to be strengthened.

As I explained earlier, from now on, I refer to forms of oppression which can be likened to (Foucauldian) ‘government of individualisation’ as ‘subjection’. I distinguish oppression in this latter sense from oppression as ‘domination’, which should be understood as including structural ‘states of domination’, radical (micro-level) power asymmetries which may leave agents incapable of “playing” the games of power, and (micro-level) ‘relationships of violence’.

### Chapter 3: Limits of Historical Models of Personal Autonomy

In Chapter 1, I justified my focus on John Christman’s historical model on the grounds that his account is promising for achieving three desiderata: (a) avoiding paternalism and perfectionism as much as possible; (b) stopping short of ascribing autonomy in implausible cases; and (c) starting a discussion on the necessary reflective materials that could make the perspectives of oppressed agents critical enough to limit their oppression. Christman’s strategy, as I explained, consists in giving agents the final say on their autonomy while ensuring that an (eventual) intellectual and affective self-assessment is well informed by the agents’ personal histories and carried out in light of the agents’ practical identities, i.e., roughly speaking, in light of the kind of persons they are.<sup>249</sup> I also noted that Christman develops his model to help elucidate when agents should be granted liberal rights and when paternalistic interventions should be limited. Christman wants to determine when agents’ preferences or values “deserve the centrality that moral and political theories place on them”.<sup>250</sup> In this chapter, I do not contest Christman’s model in light of the latter aim.

My point in this chapter is that *if* Christman’s account is to provide opportunities to challenge or (meaningfully) “endorse” oppressive heteronomous interferences that affect agents’ *developments* as a whole (like subjection does) then it is in need of further elaboration. Crucially, I defend the position that Christman’s strong reliance on sustained or settled values (especially, but not exclusively, in his most developed model) is problematic in the context that concerns me here. In a word: I argue that, when one’s development as a whole has been shaped by oppressive circumstances, commitment to one’s “authentic” values might make it

---

<sup>249</sup> ‘Eventual’, to recall, because Christman’s early model does not demand for *actual* reflection but for counterfactual reflection – i.e. *were* an agent to reflect on a trait, she would not feel alienated from it. In his revised model (discussed below) Christman speaks of a “disposition” to reflectively accept one’s motivational structure.

<sup>250</sup> Christman, “Autonomy and Personal History,” 1.

extremely difficult to take a *critical* stance vis-à-vis one's character, circumstances or developments, which should be key to carry out a meaningful procedural test.

Moreover, I argue that relying on sustained values and persistent affective reactions too much (as Christman's model does) might even "block" the possibility of experiencing 'alienation' (as defined in Christman's model) in oppressive cases which affect one's development. Since experiencing alienation is a key marker of heteronomy in Christman's procedural account, failing to secure sufficient conditions to experience alienation could make agents incapable to detect and assess interferences that they would possibly deem heteronomous.

Like Christman, I consider alienation to be a "warning sign" or an indicator of heteronomy. However, unlike Christman, I suggest that a life without deep and articulated "inner conflicts" or abrupt "short-circuits" is not necessarily a self-governed one. For example, an agent may not experience alienation because she is "numbed" to her suffering or due to an impossibility to articulate discomfort or suffering into fully fledged experiences of alienation. In these cases, one has good reasons to be at least agnostic about the results of a procedural test which relies on experiences of alienation (or the lack thereof).<sup>251</sup>

A first step towards overcoming these difficulties consists in making the *interrogation* of one's social self-conception, which appears somewhat secondary in Christman's model, a key element of critical self-assessment. Thanks to this critical attitude, I claim, agents gain more opportunities to make sense of mild or unarticulated experiences of discomfort and, therefore, to experience alienation. Moreover, I also suggest that carrying out self-critique in light of social considerations is key to avoid that the experiences of alienation that oppressed

---

<sup>251</sup> In this chapter, I deal mostly with the second and third possibilities mentioned here (i.e. with the difficulties to articulate discomfort or suffering *as* alienation). For a discussion on the possibility of being "numbed" to one's suffering, in the sense of one's suffering being "invisible" to the subject herself, see Rosie Worsdale, "Recognition, ideology, and the case of "invisible suffering," *European Journal of Philosophy* 26 (2018).

agents do feel become too psychologically burdensome. For example, the latter can happen when experiences of shame are lived as a “personal” failure.

This chapter is structured as follows:

In Section I, I argue that Christman’s twofold authenticity criterion (at play mostly, but not exclusively, in his most developed model) leaves unproblematized sustained characters and settled emotional or affective tendencies. As a result, this model cannot easily accommodate the (common) forms of subjection that I have characterised in Chapter 2. In part *b* of this section, I consider potential rejoinders to my objections in light of Christman’s notion of ‘autonomous adaptation’ to oppressive circumstances. I conclude that, even though some of my objections in part *a* need to be moderated, my main worries vis-à-vis Christman’s model remain relevant. Crucially, Christman’s model give us valuable tools to differentiate degrees of agency when agents face the same (accidental or oppressive) *disruptive* interferences, but it seems inadequate to account for autonomy in cases which involve *non-disruptive* oppressive interferences (as subjection does).

In Section II, I analyse how problematizing Christman’s notion of authenticity affects other elements of the model – i.e. the historical and competence conditions of the model. Regarding the historical conditions, I defend the idea that authenticity conditions, as they stand, encourage a “directed” reading of one’s history that lacks critical potential to challenge subjection. Drawing from the resources presented in Chapter 2, I argue the notion of authenticity gets in the way of a genealogical self-critique

With regards to competence conditions, I argue that once that we identify subjection as a common form of oppression and ‘authenticity’ as an unhelpful criterion to tackle subjection, then we should be particularly wary of defining reflective competence in a way that favours “conformist” over critical self-assessment. This is a risk if an account privileges, for example, ‘coherence’ over self-questioning through its competence conditions.

My critique of authenticity, therefore, plays a crucial role in identifying other potentially problematic features in Christman's account. Let me start by problematizing the notion of authenticity.

## I. Authenticity Conditions – a Critical Assessment

As I explained in Chapter 1, Christman's authenticity conditions play a crucial role in his model because it is through the link to the values, commitments, and even "styles" that one can call "one's own" that *self*-governed decisions and acts are identified. Those decisions or acts that are genuinely an expression of the agent's will, set limits to paternalistic interventions regardless of how "poor" or "mistaken" they may appear to someone else. Therefore, identifying these "genuine" commitments is essential to make sure that one concedes agents fundamental liberal rights *only in the right cases*, namely when by granting those rights one is not in fact safeguarding commitments that are not the agent's in the first place. In this section, I critically analyse the notion of 'authenticity' in Christman's model.

### a) 'Authenticity' - limits:

Let me now briefly recall the key elements of Christman's account of authenticity in his most developed model, *The Politics of Persons* (presented in Chapter 1). In this work, Christman defines a twofold standard of authenticity:

First, authenticity is defined negatively, as the absence of alienation from one's most basic motives for action and/or life conditions.<sup>252</sup> 'Alienation' has, for Christman, both a cognitive and an affective dimension: experiencing alienation from, for example, a character trait implies both *judging* that a character trait is not acceptable in light of who I am *and feeling*

---

<sup>252</sup> Christman, *The Politics of Persons*, 13.

that it is not acceptable.<sup>253</sup> Christman makes it clear that only *acute* negative experiences count as alienating and that, for instance, mild conflict, indecisiveness or indifference do not amount to ‘alienation’.<sup>254</sup> This qualification is meant to capture the fact that, according to Christman, ‘selves’ typically include elements which are “clearly not ideal from our point of view, but also which are not compulsions to which we are resistant”.<sup>255</sup> Therefore, to be alienated from a certain value or trait implies *strong* reflective and affective repudiation and rejection of that value or trait.<sup>256</sup>

This first aspect of the authenticity condition is meant to help detect deep incoherence within the elements that constitute one’s practical identity. To give a simple example: if one has a deep-rooted commitment to non-violence (e.g. say one self-identifies as a ‘pacifist’), then one will typically reject (i.e. feel alienated from) any violent disposition or trait of character that one may spot at any given point in one’s life. Practically, then, authenticity as non-alienation ends up giving primacy to those general (or ‘organising’) values that are a *stable* part of oneself. I justify this claim further by looking into the second dimension of Christman’s authenticity condition:

Christman also speaks about ‘authentic reflection’.<sup>257</sup> This second aspect of authenticity requires “authenticity checks” that concern the reflective procedure *itself* and guarantee that critical self-assessment is *performed* authentically. (Indeed, Christman goes as far as to argue that a pattern or style of thinking can be “ours alone”, even if it is socially influenced.)<sup>258</sup> According to Christman, one’s basic commitments are *also* part of “the *executive* function by virtue of which reflective judgment is made”.<sup>259</sup> For example, assuming minimal competency, an agent could have a tendency to be more or less thorough in her self-

---

<sup>253</sup> Ibid., 144.

<sup>254</sup> Ibid., 143-4.

<sup>255</sup> Ibid., 143.

<sup>256</sup> Ibid., 144.

<sup>257</sup> Ibid., 149.

<sup>258</sup> Ibid., 150.

<sup>259</sup> Ibid.; emphasis added.

assessments – i.e. she could consistently fall within more “impulsive” or “over-thinking” patterns – when evaluating her character traits. I say “assuming minimal competency” because this aspect of the authenticity condition should be thought as allowing for different “styles” or “patterns” of *competent* thinking. In a word, an agent has room for fashioning “her own” thinking style as long as this style is minimally rational, free from pathologies, and other elements presented in Chapters 1 and 2.

In practice, as I mentioned in Chapter 1, authentic reflection is secured via ‘*sustained critical reflection*’ (SCR), defined as “[c]ritical self-reflection repeated in a variety of contexts with similar evaluative results.”<sup>260</sup> Therefore, authentic reflection also (i.e. as it was the case for ‘authenticity’ in the first sense defined above) strives for *coherence*, albeit of a different kind. In the case of authentic reflection, coherence is necessary because the results of reflective self-assessment need to maintain throughout a variety of scenarios. Indeed, SCR is deemed trustworthy on the basis that it is a *manifestation* of the set of *stable and sustained* values that constitute who the agent is: Christman claims that SCR unveils one’s “personality”, “patterns of judgement”, and “emotional tendencies” and suggests that SCR is trustworthy *precisely* for this reason.<sup>261</sup>

It is worth noting that the two dimensions of authenticity introduced above are not unrelated. In fact, we could say that authentic reflection is a necessary condition to *trust* the absence of alienation vis-a-vis particular contents. For example, if I assess a particular feature of my character enacting my authentic “assessment style” (i.e. the way in which I “normally” perform assessments) and this assessment results (in a variety of circumstances) in a lack of alienation from this trait of my character, then I would have very good reasons to call this trait “my own”, i.e. authentic.

---

<sup>260</sup> Ibid., 152-3.

<sup>261</sup> Ibid.



At this point, we can see that what remains unquestioned by the twofold notion of authenticity in Christman's account are precisely agents' underlying sustained organizing values, deep-rooted emotional responses, long-lasting commitments, and settled character traits. Christman's authenticity condition leaves *settled* dispositions *unproblematized*.

Indeed, in a recent article, Christman acknowledges that autonomous agents typically work with *given* (social) identities and that autonomy might be achieved without the need to *problematize* (i.e. to call into question) these identities. Christman distinguishes between 4 possible levels of analysis that could shed light on what it is to act deliberately *as* a certain kind of person: Level-1 reflection considers an action merely instrumentally, i.e. "what is best to do given one's situation" (i.e. it is, basically, the application of a hypothetical imperative).<sup>262</sup> Level-2 reflection considers what one has *reason* to do "[g]iven the kind of person I *am* and the commitments I have" (i.e. Level-2 reflection typically *uses* the value framework given by one's self-conception in a social context).<sup>263</sup> Level-3 reflection questions the meaning, value, and social impact of one's identity, without necessarily calling for a separation from one's identity – i.e. according to Christman I need not ask "Should I be an X?" but can instead ask "Being an X, what is the meaning, value, and social impact of this way of being?"<sup>264</sup> Level-4 reflection implies adopting an impartial perspective, i.e. "a depersonalized viewpoint from which one asks about whether *anybody* should occupy the identity in question."<sup>265</sup> Level-4 reflection is, according to Christman, the one criticized by critics of "Kantian liberalism" for assuming a problematic "unencumbered self".<sup>266</sup>

As we can see, all 4 levels presuppose socially-defined identity categories, but levels 3 and 4 involve explicit interpretation and interrogation of one's identity in light of social

---

<sup>262</sup> Christman, "Decentered Social Selves: Interrogating Alienation in Conversation with Rahel Jaeggi," in *From Alienation to Forms of Life: The Critical Theory of Rahel Jaeggi*, ed. Amy Allen and Eduardo Mendieta (Pennsylvania: The Pennsylvania State University Press, 2018), 51.

<sup>263</sup> Ibid.

<sup>264</sup> Ibid.

<sup>265</sup> Ibid.

<sup>266</sup> Ibid.

considerations. For Christman, however, reflection which covers Levels 1 to 2 is typically enough for autonomous agency in habitual everyday contexts.<sup>267</sup>

To be sure, Christman is not suggesting that, when I act *as* a certain kind of person (e.g. in those cases in which I consider what to do *assuming* the kind of person that I am), I have no indication of whether I would *reflectively* accept my identity (e.g. as I would do if I engaged level-3 reflection). Indeed, Christman suggests that “non-conflicted” (i.e. non-alienated) action in our everyday contexts as a certain kind of person typically presupposes a reflective acceptance of one’s (social) identity. I explain how this works below:

According to Christman, acting deliberately as a certain kind of person typically gives value to one’s actions while the latter, in turn, give value to one’s identity.<sup>268</sup> Christman calls this “value-conferring” loop ‘reflexive self-affirmation’.<sup>269</sup> When this loop works well, then one consistently experiences “the recursively self-affirming affective and judgmental feedback that occurs when one acts as oneself”.<sup>270</sup> Experiences of alienation disrupt this value-conferring loop.

In Christman’s revised account of autonomy, reflexive self-affirmation indicates that one has *reasons* for acting as a certain person.<sup>271</sup> That is, *reflective* acceptance of one’s identity is still key for autonomy but, for the purposes of confirming this reflective acceptance, Christman claims, one can trust the absence of experiences of alienation in one’s everyday actions as a certain person.<sup>272</sup> Then, autonomous agents are *disposed* to “reflectively accept

---

<sup>267</sup> For example, Christman concedes autonomy to Stevens (a character from the novel *The Remains of the Day*, to which I come back in Chapter 5) even if he does not reflect “in any way more deeply than level 2”. (Ibid., 54)

<sup>268</sup> Ibid. 49.

<sup>269</sup> Ibid.

<sup>270</sup> Ibid., 54.

<sup>271</sup> Indeed, Christman claims that reflection could help “dissect and explain” one’s actions, but the *justifications* for one’s actions typically emerge from reflexive self-affirmation. (Ibid., 50)

<sup>272</sup> Ibid., 54. See also “when alienation is not an issue, acting from one’s settled identity requires no overt reflective endorsement from a detached viewpoint.” (John Christman, “Coping or Oppression,” in *Autonomy, Oppression, and Gender*, eds. Andrea Veltman and Mark Piper. (Oxford: Oxford Scholarship Online), 219.

one's values and desires in a way that is *enacted* in deliberative action".<sup>273</sup> Christman's point is then that reflective acceptance of the social meaning or value of one's identity does not necessarily need to happen through the adoption of a detached, decontextualized, disembodied, and completely external point of view vis-à-vis our social identities.

Even though I am sympathetic to the move of presenting the self-validation needed for autonomy in not-overly-intellectualistic terms, my worry (in light of my aims) is still that this solution does not sufficiently challenge *settled* characters and settled emotional tendencies. The latter might also be part of the problem when subjection is at stake.

For example, agents who have "fully" become agents of a certain social sort (i.e. who think and feel *as a* certain kind of person without any internal contradiction) in an oppressive context may be capable of functioning without major short-circuits and, therefore, they might reflexively affirm themselves in the way described by Christman. Is this sufficient for autonomy even if they never had a chance to problematize their identities in light of the oppressive context which defines and gives meaning to these identities?

Taking one's general capacity to function without short-circuits as an indicator of autonomy seems unwise given the aims of my project (i.e. securing possibilities to limit oppressive heteronomy) if we consider that the affective and reflective feedback mechanisms at play in reflexive self-affirmation may also be affected by the agent's social context. Indeed, as Christman rightfully notes, reflexive self-affirmation as a kind of person always happens through identity categories which are socially defined. That social roles are socially defined means, on the one hand, that what is included or excluded from a certain social role is socially determined – to give a classic example, being a 'wife' in certain patriarchal contexts may only be thought compatible with adopting deferential attitudes vis-à-vis one's husband. Additionally, social roles are also socially defined in the sense that these roles offer different

---

<sup>273</sup> Christman, "Decentered Social Selves," 53; emphasis added.

possible sources of social praise, honour, dignity; and (conversely) of social stigma or shame – e.g. being a “good wife” in the sense and context defined above, will imply living life in a way that is socially valued and praised.

Christman acknowledges the above when he worries that “some social forms or modes of social organization systematically cause breakdowns” in reflexive self-affirmation loops.<sup>274</sup> In other words, Christman recognises that agents might be more exposed to experiencing alienation in social situations where the “social constituents of one’s self-concept” are “unsupported” – e.g. as it may happen when the worldview of a “minority” is denigrated by the hegemonic culture.<sup>275</sup> According to Christman, cases like these show that respect for autonomy can also motivate institutional change, insofar as “[p]olitical and social institutions [...] should be designed to prevent or eradicate this sort of alienation.”<sup>276</sup>

I share the view presented above insofar as it acknowledges that experiences of alienation have not merely personal but also *social* conditions of emergence. However, my worry is that Christman’s analysis as it stands might overlook other implications of this matter: while it is true that alienation may be *increased* by particular social arrangements (as the one that Christman rightfully notes), this is only one part of the story. Additionally, as I discuss below, experiences of alienation might be *muffled* by available social meanings – e.g. I might lack social hermeneutic resources to articulate a “mild” discomfort into the “acute” rejection that Christman associates with ‘alienation’. Therefore, my point is that, for the purposes of judging whether agents live autonomously according to (social) identity categories, the lack of experiences of alienation should not be deemed authoritative (for the purposes of my model), if the conditions of possibility for experiencing alienation vis-à-vis a

---

<sup>274</sup> Ibid., 55.

<sup>275</sup> Christman, *The Politics of Persons*, 209-210.

<sup>276</sup> Ibid.

social role have not been met. I say more about how to secure these conditions of possibility in section II of this chapter.

What interests me now is to reflect on the way in which one needs to think about heteronomous interferences in general – and, indeed, about oppressive interferences in particular – for Christman’s criteria of authenticity and reflexive self-affirmation to have the potential of *limiting* these interferences. That is, what kind of heteronomous interferences or interventions are presupposed when one’s sustained values and settled emotional tendencies could provide reliable (reflective and affective) resources to limit heteronomy?

I want to suggest that, even if Christman’s model is much more sensitive to sociohistorical conditionings than the other procedural accounts which I have analysed so far, it still typically sees potentially illegitimate heteronomous interferences as *disrupting* a development which is assumed to be non-problematic. I have already argued how ‘reflexive self-affirmation’ might rely on emotional *tendencies* which are not questioned or critically assessed, and how this might be dangerous if one’s development as a whole has been shaped by oppression. Let me now briefly show how the twofold criterion of authenticity (outlined above) also works on the assumption that heteronomous interferences are typically disruptive:

I turn to Christman’s piano example (presented in Chapter 1). To recall, this example involves a person going through the memories of an abusive childhood and rejecting every disposition or character trait that could be considered a product of her time with her abuser. Christman’s point is that this agent’s piano skills are not autonomous because she enjoys them without knowing that her abuser is also responsible for these skills. However, *were* she aware of the fact that her abuser is responsible for her piano skills, she *would* feel acutely

alienated from them. Therefore, Christman claims, this agent should be deemed heteronomous vis-à-vis her piano skills.<sup>277</sup>

The agent above is able to detect a heteronomous interference (and might be able to limit the effects on her character of a relationship which she now deems abusive), I argue, because the development of the agent's history allows grasping a direction, a sense, which translates into a value framework which counters her past abuse. In Christman's example it seems to be the case that (fortunately) this hypothetical agent exited an abusive situation and developed as an agent according to values which *differ* from those which had structured (or were presupposed by) her abusive childhood. However, the agent in this example would probably have more difficulties in labelling her past "abusive" if her life unfolds in a patriarchal society and her past abuser happens to be her father or husband, regardless of whether or not the abuse is ongoing. (Think for example of the many difficulties associated with acknowledging 'marital rape' – legally, culturally, and subjectively – in social contexts where women have the "duty" to satisfy their husbands' sexual "needs".)<sup>278</sup>

To summarize then, I have argued that Christman's model as it stands does not sufficiently secure opportunities to limit subjection from the agent's perspective. In the next section, I consider if the difficulties raised so far could be overcome by turning to Christman's analysis of 'self-governing adaptation'.

---

<sup>277</sup> Ibid., 158.

<sup>278</sup> See for example Jennifer Bennice and Patricia Resick, "Marital Rape: History, Research, and Practice," *Trauma, Violence, & Abuse* 4, no. 3 (July 2003): 231.

Interestingly, Christman acknowledges this kind of difficulty in an article where he argues that a narrative conception of the self should not be taken to imply that selves *are* narratives. One of the inconveniences of fully equating selves to narratives, Christman argues, is that we cannot account for those selves struggling to articulate their stories in "their masters' tongue". Christman analyses a case of a former slave who does not have linguistic resources available to make sense of her past abuse and struggles to find them. (John Christman, "Telling Our Own Stories: Narrative Selves and Oppressive Circumstance," in *The Philosophy of Autobiography*, ed. Christopher Cowley (Chicago: The University of Chicago Press, 2015), 130) I share many of Christman's points in this article, but I believe that this kind of problem is not sufficiently integrated into his model of autonomy. Crucially, this kind of problem does not sit well with his claim that only *acute* rejection amounts to alienation insofar as, as I argue below, social meanings do not merely affect what one can *say* about oneself but also what one can experience and how one experiences it. This point will be clearer in section II in this chapter, and further developed in Chapter 5, when I analyse Stevens' example.

b) Possible rejoinders:

In the remainder of Section 1, I turn to Christman's reflections on "self-governing adaptation".<sup>279</sup> In *Coping or Oppression*, Christman grants (in certain conditions that I explain below) some degree of autonomy to agents who *change* "key aspects" of their practical identities in reaction to external circumstances.<sup>280</sup> This could present an obvious challenge to my previous assessment of Christman's model because, in cases of self-governing adaptation, it might be possible to change (and perhaps to question) settled characters and sustained values. However, as I show below, the core assumptions that motivated my critique of Christman's model still hold.

To unpack and justify my latter claim, I start by briefly summarizing the study cases on which Christman's reflections are based. Christman presents four hypothetical biographies: two of them (Bernice's and Abby's) describe the lives of agents after an accident that leaves them paralysed and the other two (Kaew's and Irina's) narrate the stories of two women who (after being deceived) become victims of sex-trafficking. As we can see, while the two cases which involve sex-trafficking describe a situation that clearly strikes as illegitimate, nothing in the two cases which involve accidents triggers the same intuition. Bernice and Abby's life stories are marked by unfortunate circumstances but not (*prima facie*, at least) by oppression.

From each pair of agents who share external circumstances (i.e. two women suffered an accident and two women were victims of sex trafficking), Christman presents one of them as adapting to her new circumstances and one of them as resisting adaptation (i.e. as holding on to previous values and life plans). Regarding the two cases which involve an accident that results in permanent disability, Christman describes Bernice as being unable to adapt to her

---

<sup>279</sup> Christman, "Coping or Oppression," 201; emphasis added.

<sup>280</sup> *Ibid.*, 206.

new situation (she becomes an alcoholic, becomes isolated, and lives a life “consumed by resentment”)<sup>281</sup> while Abby, originally an athletic girl fond of marathons, picks a “long-abandoned interest in writing” and becomes a successful writer.<sup>282</sup> When it comes to the two women who are victims of sex-trafficking, Christman explains how Kaew chose to stay with her abusers once she managed to pay back the money she “owed” them and eventually became the manager of the “bar” in which she worked.<sup>283</sup> Irina, however, never ceased to hate her exploiters and never stopped making elaborate plans to escape, even when she repeatedly failed to do so.<sup>284</sup>

How could we account for autonomy in these cases? One of Christman’s points is that an exclusively “structural or external” analysis of autonomy (i.e. one that merely focuses on “sociorelational settings” to attribute autonomy) fails to explain why we may attribute *different* degrees of autonomy to agents who *share* external conditions.<sup>285</sup> According to Christman, the asymmetries in the readers’ “intuitive” reactions to similar cases can only be explained if agent’s subjective strategies are also taken into account – i.e. it is relevant to consider how agents “cope” with their circumstances. In a word, according to Christman these four examples show the advantages of *procedural* over substantive models of autonomy – I discuss the latter models in Chapter 4.

I want to suggest that, when we consider Christman’s analysis of the examples in *Coping or Oppression* which involve oppressive events, adaptation is not so clearly presented as a self-governing strategy. For example, according to Christman, common or expected reactions to Irina’s and Kaew’s cases are as follows:

---

<sup>281</sup> Ibid., 204.

<sup>282</sup> Ibid., 202.

<sup>283</sup> Ibid.

<sup>284</sup> Ibid., 204. Christman presents a more detailed account of these four cases. I have simplified these descriptions for the sake of brevity.

<sup>285</sup> Ibid., 208.



At first glance, at least, it seems that that Irina retains a modicum of self-government just because she resists and resents her constraining condition, while Kaew might be seen as losing her autonomy just because she has adapted to her condition and altered her identity in response to it.<sup>286</sup>

Christman suggests that one would “intuitively” grant Irina some degree of self-government because, by hating her oppressors and keeping her desire to escape, she shows *resistance* to her oppression. Resistance is then an indicator of (some degree of) self-government and the former is pictured as the preservation of a form of self-understanding that is not *merely* the product of her oppressive circumstances – i.e. some of Irina’s values and life plans continue to be those she had *before* becoming a victim of sex-trafficking.

Admittedly, the “intuitive” reaction to Irina’s case that Christman appeals to is not an easily sharable one. In fact, both from a procedural and from a substantive point of view, it could be argued the opposite, i.e. that Irina *lacks* self-government. For example, from a procedural perspective à la Mele, one may argue that this whole situation implied a bypassing of Irina’s rational capacities (according to Christman’s description, she ended up in her present situation because she was *deceived* by sex-traffickers),<sup>287</sup> and that the effects of that bypassing typically stay in force in cases like these (e.g. if agents “choose” to stay with their oppressors while under threat). In other words, one could question if the external conditions that are typically a part of some historical procedural models satisfy in Irina’s case and, therefore, if she could qualify as autonomous.

Furthermore, the intuitions of those defending some form of substantive approach to autonomy would clearly identify Irina’s case as one in which self-government should be ruled

---

<sup>286</sup> Ibid., 204-5

<sup>287</sup> Ibid., 201.

out because she is *de facto* non-autonomous.<sup>288</sup> Irina’s strategies might be admirable in many ways but the fact remains that she should have never been put in this position in the first place. I say more about substantive accounts in Chapter 4.

Similarly, in my own position (developed in Chapter 5), Irina’s case could be seen as one involving domination and, as such, external changes would be needed before Irina’s capacity to limit the effects of her oppressive conditions could be assessed procedurally. We can indeed recognise her as resisting, but resisting – I argue in Chapter 5 – falls short of autonomy. But let me leave this potential disagreement on our “intuitive” reactions to Irina’s case aside for now and go back to considering how Christman pictures oppression in his examples.

A similar strategy (i.e. valuing an agent’s capacity to hold on to values or commitments held *before* oppression happened) applies to other cases involving oppression discussed in *Coping or Oppression*. Consider for instance Christman’s analysis of ‘resistant slaves’, i.e. slaves who “never cave in to the oppressiveness of their conditions”.<sup>289</sup> Christman argues:

In all such cases, two important factors remain in play. First, these agents maintain a practical identity that is continuous with their *pre- or nonoppression selves*. Their value priorities and senses of themselves are not crushed by their captivity, even if they are prevented from acting on them. [...] I am not claiming that such people are in any way fully autonomous, [...] But I do think there is another sense of minimal autonomy that they do maintain, one that importantly distinguishes them from those who are crushed by circumstance into either internalizing the oppressive values structuring their

---

<sup>288</sup> See for example Marina Oshana, *Personal Autonomy in Society* (Aldershot: Ashgate, 2006), 2. More on substantive accounts in Chapter 4.

<sup>289</sup> Christman, “Coping or Oppression,” 210.

domination or losing completely any effective power to act on their own at all.<sup>290</sup>

Christman's identification of (some degree of) autonomy under oppressive circumstances relies again on the existence of a self that is not reducible to oppression, i.e. this self is not *merely* an effect of oppressive circumstances. And how could we have access to a self not entirely defined by existing oppressive circumstances?

The first way in which one could identify a "non-oppressed self" is via reference to values and motives that were "one's own" *before* oppression started – as Christman puts it, one could have access to a "*pre-oppressed self*". This way of framing the issue, however, would only work when one can clearly identify a "before and after" oppression. As I have already suggested, common forms of oppression like gender oppression do not admit this temporal demarcation. Indeed, as Sandra Lee Bartky claims, gender oppression *differs* from other forms of oppression (e.g. colonialism) precisely because of the impossibility to think of a "time before" oppression. As Bartky puts it: "women have no memory of a 'time before': a time before the masters came, a time before we were subjugated and ruled".<sup>291</sup>

Someone could perhaps argue that even though 'women' as a collective could not have a memory of a time before their oppression, individual women could identify concrete moments in their biographies when they *became* affected by oppressive gender norms in a way they were not before. For example, the case could be made that individuals might come across particularly oppressive aspects of gender at different moments in their lives (e.g. when one is considering different careers or life paths and certain options appear closed because of one's gender). If this were an accurate representation of gender oppression, then an agent could perhaps think of different moments of her life as happening "before and after"

---

<sup>290</sup> Ibid.; emphasis added.

<sup>291</sup> Bartky, "On Psychological Oppression," 25.

becoming affected by oppressive gender norms and might even have access to values and commitments that she could call “pre-oppressive”.

It might be true that certain aspects of gender oppression may become more pressing or more noticeable at specific moments in one’s life. Cases like the ones mentioned just above (i.e. agents realising at a specific point of their lives how certain life options are severely restricted because of their genders) could be considered examples of this. However, I believe that it would be a mistake to think that these agents had not been subjected to gender norms before the moment in which their force becomes noticeable. Developing as a ‘female’ or as a ‘male’ subject entails a long list of lifestyles and characters that one could (or could not) pursue and these possibilities and impossibilities exist “from the beginning” regardless of whether one experiences these injunctions or limits first-hand.<sup>292</sup> Indeed, if a specific norm concerning femininity or masculinity has a force on oneself in the first place it is because it touches on crucial aspects of who one is or considers oneself to be and this self-conception as a gendered being could be hardly considered as being free from problematic gender injunctions. Therefore, referring to values and commitments that I had before being “consciously” constrained by particular aspects of my gender will not point in the direction of an identity untouched by gender norms.

Another potential rejoinder is that a “non-oppressed self” could be found *even* in cases where oppression does *not* allow for a division between “pre-oppressive” and “post-oppressive” times. For example, the phenomenon of ‘double-consciousness’ may still provide a historical account like Christman’s with a “non-oppressed self” that could make the model work even under non-disruptive oppressive circumstances, e.g. in cases involving

---

<sup>292</sup> See Allen, *The Politics of Our Selves*, 165-8 for a discussion on the empirical literature in developmental psychology on gender development in infants. The consensus is that gender becomes salient at least as early as the age of two.

living one's whole life subjected to racist stereotypes.<sup>293</sup> In a scenario allowing for the development of 'double-consciousness', a subject who has developed while exposed to a degrading self-image available in the dominant culture could still find positive forms of self-definition thanks to the solidarity among people who share her situation. In such a case then, an agent would have access to a "non-oppressed self" regardless of the fact that oppression could not be deemed "disruptive".

In reply, let me say, first, that the assumption that the alternative self-image that the oppressed may be able to maintain or form will be unaffected or undistorted by the dominant culture seems to me hard to accept. That is, even when the oppressed could have access to two coexisting "self-conceptions", the one that forms in opposition or in reaction to the dominant culture will not necessarily be free from the normative evaluations of the dominant culture. An example of this could be the phenomenon that Gustavo Pereira calls 'malinchismo,' namely a social pathology particularly present in Latin America which he traces back to the colonial past of Latin American societies. Pereira defines 'malinchismo' as "a distorted relationship with oneself and with others, by which the culture in which someone was born and grew up is considered of a lower value than foreign ones" (i.e. than the cultures of one's –former or past– colonisers).<sup>294</sup>

Second, it seems that the 'double-consciousness' phenomenon somehow presupposes that those oppressed are, at least, (intellectually and/or affectively) aware of the social

---

<sup>293</sup> Iris Marion Young considers the phenomenon of "double consciousness" when analysing the psychological effects of cultural imperialism (one of the "faces of oppression" that she proposes). According to Young, double consciousness "arises when the oppressed subject refuses to coincide with these devalued, objectified, stereotyped visions of herself or himself." (Young, *Justice and the Politics of Difference*, 60) Young describes a duality between a negative self-image imposed by the dominant culture and a positive one that survives or "emerges" in spite of that dominant culture. For Young, the collective nature of oppression makes this double consciousness possible, for oppressed people do not only suffer from forms of misrecognition but also can benefit from forms of solidarity. Because the culturally imperialized can "affirm and recognize one another as sharing similar experiences and perspectives on social life", they can "often maintain a sense of positive subjectivity." (Ibid.)

<sup>294</sup> Gustavo Pereira, *Imposed Rationality and Besieged Imagination: Practical Life and Social Pathologies* (Cham: Springer, 2019), 152.

*contingency* of the self-image that they contest, which would not be the case when one's development, and the self-images and values derived from this development, are merely deemed "normal" or, even, "true". In this sense, I share Allen's worry, when she expresses that the double consciousness model "does not take seriously enough the ways in which [some] oppressed individuals often more or less *fully* accept and *identify* with the negative cultural images of their own group as the basis of their own self-image".<sup>295</sup>

The case of gender is once again helpful to illustrate my point: processes of subject formation according to gender norms happen according to norms that are "internal" to *one's* culture and, therefore, gender oppression does not so easily allow for this "us and them" logic that may be the case in, e.g., colonialism. Bartky claims:

Unlike the black colonial [...], women qua women are not now in possession of an alternate culture, a "native" culture which, even if regarded by everyone, including ourselves, as decidedly inferior to the dominant culture, we could at least recognize as our own. However degraded or distorted an image of ourselves we see reflected in the patriarchal culture, the culture of our men is still our culture.<sup>296</sup>

In short, the formation of a "non-oppressed" gendered self that could develop together with a gendered one is neither a necessary nor straightforward occurrence.

What needs to be retained is that, in the majority of the cases proposed by Christman which involve oppression, the criterion to distinguish those agents who keep some degree of autonomy from those who do not, seems to be given mainly through the commitment to a self who is *not* the product of the oppressive circumstances which are trying to be limited.

---

<sup>295</sup> Amy Allen, "Power and the Politics of Difference: Oppression, Empowerment, and Transnational Justice," *Hypatia* 23, no. 3 (July-September 2008): 162; emphases added.

<sup>296</sup> Bartky, "On Psychological Oppression," 25.

Therefore, my worry that Christman's model could not deal with cases of oppression, when by the latter we mean forms of 'subjection,' still holds, because subjection does not allow for the development of an "oppression-independent" self.

Before concluding this section one additional point is necessary. Even if, as I argued above, Christman's analysis of self-government under oppression in *Coping or Oppression* relies on the possibility of relating to a "non-oppressed" self, Christman does consider cases when people's identities *change* quite radically and their autonomy *maintains*. Abby's case, for example, shows that autonomy is compatible with changes which result in a "new way of looking at and experiencing the world".<sup>297</sup> Does the latter possibility imply a complete departure from the conditions that regulated autonomy in Christman's early model (i.e. commitment to one's settled character and sustained values)?

Upon closer examination, however, Abby's case does not seem to imply a radical departure from her settled identity before the accident. Abby's case could be better described as a "reordering" of values than as a radical change of self-identity. Abby seems to *maintain* the values that she had before the accident and what changes is the agent's value-priority or the degree to which different values are (and could be) realized. In other words, it seems that Abby reshapes her life according to values that were, in a sense, already *hers* before the change of circumstances.<sup>298</sup>

Still, Christman does consider processes of self-governing adaptation that are more radical and which seem to imply a bigger departure from previously-held commitments or values. This is the case when Christman wonders if Kaew's case could possibly be one of 'self-governing adaptation'. To recall, Kaew's adaptation involves (what appears to be) a departure from previous values and a change of perspective: she embraces her role as a sex-

---

<sup>297</sup> Christman, "Coping or Oppression," 207.

<sup>298</sup> Christman speaks about Abby turning to a "long-abandoned interest in writing", which suggests that this interest does not come out of nowhere. (Ibid., 202)

worker when her “debt” is paid and stays in the place she once wanted to escape. Christman would deem Kaew autonomous under certain conditions:

Is a person like Kaew, for example, best seen as one who embraces the identity of a resilient survivor and who engages in prostitution because her life options limit her to such choices as a method of, say, providing for her family? If so, then if other conditions of autonomy are met we should not withhold that label to her.<sup>299</sup>

Christman admits that cases like Kaew’s are quite rare and that they would be typically incompatible with reflective competence conditions.<sup>300</sup> Still, if other autonomy conditions satisfy, Christman would be willing to deem an agent like Kaew autonomous as long as she “successfully” relates to an available identity category. The latter means, as I have explained above, that Kaew should be capable of reflexive self-affirmation.

Therefore, a key point in *Coping or Oppression* is that reflexive self-affirmation does not always need to happen in light of the same settled identity category: as Abby’s and Kaew’s cases shows, a same individual might function in a reflexively self-affirming way thanks to different practical identity categories throughout her life (e.g. ‘marathoner’, ‘writer’, ‘sex-worker’, ‘survivor’).

Could this solution help overcome the difficulties with the notion of authenticity that I suggested before? Does acknowledging that reflexive self-affirmation could happen through different identity categories avoid the problematic commitment to one’s sustained values which I signalled earlier? Admittedly, this solution accommodates change in autonomous agents’ value commitments better than Christman’s early model did. However, I argue, as it stands, this criterion is not necessarily better suited to limit subjection:

---

<sup>299</sup> Ibid., 224.

<sup>300</sup> Ibid., 223.



First, let me note briefly that, in the examples presented above, oppression continues to be seen as something that “strikes” an individual’s life at a particular point in time. As I have argued earlier, picturing oppression in this way is unhelpful to think of strategies to limit subjection. Importantly, in the cases described above, the change in individuals’ practical identities, self-relations, and emotional reactions is motivated by an abrupt and noticeable change of circumstances. To tackle subjection, however, the key question is how we could explain and even encourage revision and change of one’s perspective and settled commitments *even* when agents’ external circumstances stay the same.

Second, even if I argue that to limit subjection we need a model which enables agents to question (and, possibly, to change) sustained commitments and settled characters, I do not imply that *any* change of perspective or value framework will necessarily make agents more autonomous. As I have suggested before, I am interested in a model of autonomy that enables agents to be *less* governed by others, not merely to be governed differently. Therefore, what should remain central is that agents relate to (old or new) practical identities while at least minimally aware of the *social* implications of these practical identities. For subjection to be limited, I argue, it is key that agents see themselves and their (conforming, transgressive or coping) actions in light of their *social* circumstances.

Indeed, managing to avoid experiences of alienation by embracing a different practical identity is not necessarily a step away from heteronomy. If, for example, a “new” practical identity is similarly functional to the same oppressive structure as the “old” one was and, moreover, if this functionality remains unseen or unproblematized by the individual concerned, then adaptation would not limit oppression. I illustrate my point by going back to Kaew’s case:

What sort of self-assessment would be necessary for Kaew to limit oppression? Let me note that (as I suggested earlier and as I discuss in Chapter 4), since Kaew’s margin for action seems to be so severely reduced by a wide variety of external constraints (e.g. her race

and immigration status in a foreign country, her dependence on her exploiters, her lack of access to alternatives, etc.), it is more appropriate to think of Kaew as dominated than as subjected. That is, considerable external change appears necessary in Kaew's case before her subjective strategies could significantly limit her oppression.

Allow me, nonetheless, for the sake of the argument, to assume that Kaew is "just" subjected. Imagine, for example, that Kaew does not live under threat, that she is minimally protected by the law, that she freely conducts herself *as* a 'woman', member of a certain ethnic minority, individual of a certain class, etc. My point is then that Kaew could be similarly subjected by virtue of her gender, race or class even after adapting and managing to function without "short-circuits" through a different identity category. To be sure, the category of 'resilient survivor' may make her situation more bearable in important ways (e.g. by reducing suffering or anger or by improving her self-esteem) but it does not necessarily make Kaew less oppressed.<sup>301</sup> For example, if 'resilient survivor' is just a different placeholder for the same role in the same oppressive structure and Kaew relates to this new category without at least a minimal understanding of the social role that she *continues* to play then, I argue, oppressive heteronomy is not significantly limited.

Additionally, to use the language presented in Chapter 2 (when analysing genealogy), Kaew needs to be minimally aware of what different social identities and practices *do*, of their social *purposiveness*, of who they benefit and who they burden. Imagine (to adapt an example provided by Srinivasan) that Kaew rejects the term 'prostitute' and relates to the identity of 'resilient survivor' as a way of, say, exposing "women's submissiveness as a fiction".<sup>302</sup> Moreover, she even sees sex-work as a way to practically subvert this social injunction to submissiveness – e.g. she judges that sex-work allows her to be in control of her life

---

<sup>301</sup> I distinguish between 'resisting' (e.g. as avoidance of annihilation) and autonomy in contexts of domination in Chapter 5.

<sup>302</sup> Srinivasan analyses the displacement of 'prostitute' with 'sex-worker'. (Srinivasan, "Genealogy, Epistemology, and World Making," 145).

financially and practically in a way that she prefers over other options actually opened to her (given also other social circumstances such as her race or class). Then, self-affirmation through the category of ‘resilient survivor’ or ‘sex worker’, may indeed be a strategy to limit (some) aspects of her subjection in relevant and significant ways.<sup>303</sup>

Note, however, that the crucial question to assess whether Kaew’s strategies manage to limit her oppression is not merely whether she managed to avoid alienation and function without short-circuits. Rather, it is essential to ask whether the kind of self-assessment carried out secures the cognitive and affective conditions of possibility for experiencing alienation when subjected and, when applicable, for developing strategies of self-transformation aimed at limiting oppression. I analyse further how a more *social* self-understanding is crucial for experiencing alienation when subjected in the next section.

## II. Consequences of a Critique of Authenticity over Other Elements of the Model

What are the effects of challenging the criterion of ‘authenticity’ over the other elements of Christman’s historical model? In what follows, I analyse these effects in turns:

### a) Problematizing Christman’s historical condition:

In this section, I consider the effects that problematizing authenticity conditions has on the historical condition of Christman’s model. How are individuals to relate to their histories for critical self-assessment to be possible and, crucially, for this self-assessment to be strong enough to challenge settled ways of being and feeling? I argue that, to answer this question, a historical model of autonomy should reflect on at least two issues: First, it is

---

<sup>303</sup> I want to thank Amelia Horgan for a very thought-provoking discussion on sex work which, as I see it now, helped me think about this example.

necessary to consider if different “readings” of one’s history may affect one’s capacity to relate critically to one’s settled values and traits of character. Second, it is necessary to reflect on whether all histories, no matter their scope, could contribute to a critical self-assessment capable of challenging one’s subjection.

I consider the two abovementioned issues in light of two key lessons drawn from my analysis of genealogy in Chapter 2:

First, since we are working with *historical* materials, we explicitly need to address the fact that different readings and interpretations of the same histories are possible and that not all of them might bring about the same critical results. For example, it is key to reflect on how different possible readings and interpretations of one’s history may be conditioned or directed by the “meta-histories” guiding our history-telling. Crucially, I argue that the ideal of authenticity presented above might promote an unhelpful interpretative strategy if one has in mind the purposes of limiting subjection – namely authenticity conditions may favour conservative over critical readings of one’s history. What I mean by this is that, if historical self-assessment is carried out assuming that settled traits of character are “safe” indicators of who one “really” is, then critical self-assessment might be approached more as a mode of self-*discovery* in light of one’s personal history than as a matter of self-*questioning* in light of historical social determinations. To limit subjection, however, the former strategy will not take us far enough.

And what kind of historical materials are more suitable to promote forms of critical self-assessment capable of limiting subjection? As I anticipated, I argue that effective opportunities for testing settled values and emotional tendencies are increased when agents look beyond their *personal* histories and, crucially, when (as a result of this change of historical “scope”) they stop interpreting some traits as merely “personal” features. As I suggested in Chapter 2, for the purposes that concern me here, agents should assess values and

commitments in light of a history that is not merely personal but also that of their contemporaries and, more specifically, that of those in similar social situations.

The above means that agents need to be minimally concerned by social roles and by the histories and perspectives of others around them. Consider for example how knowing about others who occupy similar positions in society, or who go through the same everyday struggles, might be crucial to increase one's self-understanding *qua* 'woman', 'man', 'Black person', etc. Becoming aware of the fact that an always assumed "character" trait is actually the unquestioned acceptance of a socially available model for being a "normal x" (woman, Black woman, man, etc.) might only be possible after I connect my "personality" with those of others who occupy similar social positions.

To give a simple example for now: say that I want to assess a certain well-established character trait (e.g. my being "shy" or "introvert") in light of its history. If I proceed assuming that this feature is a part of who I *really* am or taking for granted that the trait is just a *personal* trait, I am likely to focus on events in my personal history and to use elements or episodes in my biography to *explain* this trait. Say that instead of proceeding in this way, I am suspicious of the fact that most women in my circles happen to be "introverts" – e.g. I may be puzzled by the fact that many of my women colleagues also happen to struggle to voice their opinions in public. This change of "meta-narrative" will most likely affect the way in which I read the stages in my development leading to this particular feature – e.g. I may now see how the enactment of gender roles during family dinners or at school might have contributed to shape my character. Additionally, "de-individualizing" this character trait will certainly have an effect on which historical information I consider relevant to assess the origins of this trait. For example, I might now see crucial to also pay more attention to the histories of other women.

Moreover, my suggestion is also that linking a trait of character with the social injunctions which prescribe how *people like me* should be (i.e. with the sociohistorical spaces

of possibility for character formation), affects agents' capacities to experience alienation.<sup>304</sup> In the remainder of this section I explore the connection between the scope of the history which informs one's self-assessment and experiences of alienation.

A second genealogical point to recall from Chapter 2 is that different forms of history-telling have different intellectual and affective effects on those whose histories are being told. Since I analyse here the conditions of possibility for experiencing alienation, I am particularly interested in the *affective* effects which different historical self-assessments could have on agents. Placing one's personal history within a wider social history and assessing one's character in light of the spaces of possibility for character formation available in one's social context, I argue, may enable agents to relate differently to their affective experiences. I unpack the different senses in which this might be possible below:

First, I refer to my earlier point about the possibility of making sense of negative experiences (e.g. of experiences of abuse or discrimination) thanks to *collective* hermeneutical work. Miranda Fricker's analysis of consciousness raising feminist groups supports this claim. Fricker argues:

If we look at the history of the women's movement, we see that the method of consciousness raising through 'speak-outs' and the sharing of scantily understood, barely articulate experiences was a direct response to the fact that so much of women's experience was obscure, even unspeakable, for the isolated individual, whereas the

---

<sup>304</sup> The introduction of the problem of "social roles" in the context of a discussion on the historical materials necessary for autonomy might seem odd to some readers. However, it is worth recalling that the critical self-problematism presented in Chapter 2 was described as a "historical ontology of ourselves". In this sense, historical determinations do not merely include one's past, nor even just the pasts of "people like me". Additionally, one is historically determined in the sense that one's possibilities for living and acting are limited and enabled by the spaces of possibilities opened by socially available roles in a specific sociohistorical context. As Ian Hacking summarizes it: "Historical ontology is not so much about the formation of character as about the space of possibilities for character formation that surround a person, and create the potentials for 'individual experience.'" Ian Hacking, *Historical Ontology* (Cambridge: Harvard University Press, 2002), 23.

process of sharing these half-formed understandings awakened hitherto dormant resources for social meaning that brought clarity, cognitive confidence, and increased communicative facility.<sup>305</sup>

The phenomenon described by Fricker is especially relevant if we consider the importance that experiencing alienation has in Christman's model of autonomy. Indeed, a model of personal autonomy that relies on experiences of alienation to decide autonomy, should avoid perspectives that do not secure the conditions of possibility for experiencing alienation. If some negative experiences cannot be fully grasped or articulated until they are, for example, shared by those in similar social situations, then considering the perspectives of others might increase opportunities for testing one's general values and dispositions through a procedural test.

Therefore, a first sense in which a procedural self-assessment might benefit from a broader social perspective is by facilitating the articulation or intensification of "mild" or "confusing" negative experiences which may be present but not experienced as 'alienation' (in Christman's sense). This is, however, only one part of what I want to argue.

Additionally, carrying out self-assessment with at least a minimal awareness of one's social situation is promising because it enables a different relation vis-à-vis obvious or strong affective experiences. Being "numbed" to one's suffering or being confused by one's discomfort may be frequent occurrences but so are experiences of guilt, shame, frustration, fear, and anger all too common in the lives of the oppressed. In this latter case, a change from a personal to a social historical perspective may unburden agents by showing the social nature of their suffering. For example, instead of experiencing shame or guilt as a personal failure, agents might question the expectations associated with the social roles which trigger these negative experiences. In order to justify this point, allow me a brief detour – I consider

---

<sup>305</sup> Miranda Fricker, *Epistemic Injustice: Power and the Ethics of Knowing* (Oxford: Oxford Scholarship Online, 2007), 148.

below the effects that different ways of inhabiting social roles could have on individuals' regard or disregard of their own experiences while playing these social roles.

According to Jörg Schaub, when we analyse different forms of “inhabiting” social roles and of “meeting [the] behavioural expectations associated with them”<sup>306</sup>, it is possible to distinguish between “conformist” and “conforming” attitudes vis-à-vis the same social roles. Significantly, when conformism is the case, individuals *cease* to guide their enactment of different roles by how they *experience* playing those roles.<sup>307</sup>

Schaub draws from the sociological literature on role theory and argues (with Honneth) that “role-mediated relationships engender a ‘tendency for conformism’”.<sup>308</sup> This tendency is explained by sociologists in several ways – e.g. individuals typically attempt to maintain social roles (and the relationships enabled by these roles) “stable” and “predictable”; and individuals “attach” to their roles and therefore “merge self with given roles”.<sup>309</sup> According to Schaub, the stronger the attachment to our roles, the stronger the tendency to conformism will be and “the more inclined we become to *disregard* experiences with our roles that bring us at odds with them.”<sup>310</sup>

To be sure, attachment and conformism do not come out of nowhere, as Schaub notes these phenomena are explainable by the fact that different social roles are a source of social esteem. Therefore, individuals’ “craving[s] for favourable evaluations” encourage them to play roles in ways which they see as being more conducive to social praise.<sup>311</sup> Crucially,

---

<sup>306</sup> Jörg Schaub, “Aesthetic Freedom and Democratic Ethical Life: A Hegelian Account of the Relationship Between Aesthetics and Democratic Politics,” *European Journal of Philosophy* (2018), 11.

<sup>307</sup> Ibid.

<sup>308</sup> Ibid., 12

<sup>309</sup> Turner in Ibid., 13.

<sup>310</sup> Ibid.; emphasis added.

<sup>311</sup> Ibid., 13. The argument that subjects attach to their social identities, even when the latter are oppressive or a source of suffering is not uncommon. Judith Butler argues (via a psychoanalytic reading of Foucault) that the formation of the psychic and social identity of a child requires attachment to those on whom the child is dependent. The desire for one’s social existence and persistence, Butler argues, leads to a desire for one’s own subordination. Butler claims: “Although the dependency of the child is not political subordination in any usual sense, the formation of primary passion in dependency renders the child vulnerable to subordination and exploitation, ...” (Judith Butler, *The Psychic Life of Power: Theories in Subjection* (Stanford: Stanford University Press, 1997), 7-9.



the “craving” for favourable evaluations can even lead to agents “acting” their experiences of social roles or to individuals “training” themselves to experience roles as it is socially expected. This is so because feeling differently from what it is socially expected one should feel when enacting a role can be seen (by others) as a bad enactment of a role and felt (by oneself) as a personal failure.<sup>312</sup> Consider, for example, how feeling dissatisfied with or overburdened by one’s role as an ‘entrepreneur’ might be seen as being at odds with the social expectations of this role and, indeed, with the kind of satisfaction that this role is *supposed* to provide.

Schaub’s analysis of conformism allows us to identify a “natural” (or, at least, a historically-encouraged) tendency which complicates matters for procedural accounts like Christman’s: when we enact roles in a conformist way “we do not let ourselves be guided by how we *experience* playing roles.”<sup>313</sup> Conformism, then, could lead to a form of alienation more “fundamental” than that which is the focus of Christman’s analyses insofar as conformism could block the emergence of experiences of alienation à la Christman. Significantly, negative experiences (regardless of their intensity) will not necessarily trigger “breakdowns” in one’s enactment of a role simply because agents may prioritise social expectations over their own experiences. Indeed, negative feelings, no matter how acute, may be just overlooked or interpreted as evidence that one is not acting a role well enough instead of leading agents to question, for example, whether a social role undermines their autonomy. Moreover, recurrent negative experiences may be simply decoded as “personal faults” – e.g. as evidence that, no matter how hard one tries to conform, one is simply not good enough for a role that one wants to inhabit or perform well.

In other words, experiencing ‘alienation’ in Christman’s sense (e.g. experiencing that there may be a “mismatch” between who one “really” wants to be and the practical identity

---

<sup>312</sup> Schaub, “Aesthetic Freedom and Democratic Ethical Life,” 14.

<sup>313</sup> *Ibid.*, 13; emphasis added.

which one enacts) requires, at least, that one *lets* oneself be guided by one's experience of a role. The latter precondition, as I argued above, cannot be taken for granted.

My suggestion is that the change of scope which I have been describing (from one's personal history and perspective to one's social history and perspective) may contribute to secure the preconditions for experiencing alienation (in Christman's sense). Relating to one's experiences in light of, for example, the social expectations of the roles which one inhabits helps in at least two ways:

First, by unveiling the social character of certain ways of being (e.g. by showing that a personality trait is not "who I really am" but a trait embedded in a social role) the "merging" of self and social role that Schaub mentions as conducive to conformism can be pushed back. To be sure, the "self" (and the experiences) that I could access by taking some critical distance from the social roles I happen to inhabit, is not one that exists independently of any social role or category. Inferring the latter would be at odds with the Foucauldian picture which I have presented. Instead, my suggestion is that adopting a social perspective is key for problematizing settled traits of character and affective reactions more openly *as* the product of social injunctions. This includes gaining awareness of the continuous (affective and cognitive) "training" of ourselves that playing social roles implies.

Second, a change of scope towards a more social history and perspective may also make suffering less "personal". For instance, agents may be able to see that experiences of shame or guilt are connected to the excessive or inadequate expectations of a social role and are not "their fault". This is a first step towards avoiding that suffering erodes the minimal self-esteem needed for challenging social expectations and, therefore, for having opportunities for limiting heteronomy under oppressive circumstances.

The two points above are underdeveloped for now but become clearer through the analysis and examples presented in chapters 4 and 5.

In the remainder of this chapter, I turn to the third condition set by Christman's model. I consider how 'reflective competence' needs to be understood once we challenge the view that "authentic" values need to be preserved and that history should be read with a view to discovering and explaining those authentic values.

b) Problematizing Christman's competence conditions:

My aim in this section is to signal the limits of Christman's competence conditions in light of the problematisation of authenticity presented above. As I suggested earlier, a procedural model of personal autonomy that wishes to take subjection seriously should be particularly wary of privileging "self-discovery" over self-questioning or critical self-assessment. This has implications for competence conditions too, namely reflective competence cannot be defined in such a way that coherence and continuity of one's sustained commitments are valued over self-questioning.

In other words, as I argued in Chapter 2, the capacities necessary to carry out the self-assessment involved in autonomy should be defined having in mind the kinds of interferences or interventions which an autonomous subject should be able to assess and, possibly, to resist or limit. Once that we have challenged deception and manipulation as the paradigmatic models of heteronomous interference, the repertoire of reflective competences necessary to critically assess one's situation needs to be understood differently. For example, being capable of maintaining a coherent set of organising values might not necessarily be a helpful intellectual skill to tackle subjection – i.e. if agents also need to challenge settled dispositions and stable values. In this context, being able to step back from forms of self-management that happen according to "normally" acquired guidelines or, even, being able to imagine oneself otherwise might become much more relevant skills.

I start by briefly recalling Christman's characterisation of reflective competence in his most developed model. Christman defines both 'reflective competence' and 'competent agency'. I focus especially on the former definition because Christman's characterisation of a competent agent partly relies on an understanding of what competent reflection is – i.e. a competent agent is able to *think competently*, and she is also able to *direct her conduct* in light of competent reflection. In other words, a competent agent is one who puts her reflection into practice and lives and acts in a way that mirrors what she reflectively decided. Christman explains:

...a competent agent is minimally rational (where, for example, her desires and plans contain no manifest contradictions that could be easily brought to consciousness), she displays minimal self-control, and is generally able to form effective intentions that in the absence of external barriers lead to completed action.<sup>314</sup>

As we can see, then, apart from the claim that reflection needs to be minimally coherent, this definition leaves unproblematized what kinds of reflective processes should "inform" one's actions. So, this definition is not very useful to understand what constitutes "good" reflection for the purposes of autonomy. For example, this picture of competent agency could be compatible with coherent delusions or even with some mental disorders (e.g. people suffering from anorexia nervosa typically exhibit very coherent desires and plans and extreme forms of self-control).<sup>315</sup>

What this fragment does show, however, is that the kind of reflection which is relevant for autonomy is not valued by itself. Competent deliberation needs to be effectively connected with one's capacities to form motivations for action (on oneself, on the world,

---

<sup>314</sup> Christman, *The politics of persons*, 155.

<sup>315</sup> See for example Fabian Freyenhagen and Tom O'Shea, "Hidden substance: mental disorder as a challenge to normatively neutral accounts of autonomy," *International Journal of Law in Context* 9, Special Issue 01 (March 2013): 61.

etc.). A competent agent does not merely reason in an autonomous way (which implies a reflective competence condition) but also succeeds in making this reflection inform and direct her conduct.

Christman claims that for reflection to be meaningful or trustworthy, it needs to happen under certain conditions: it needs to be “both minimally competent and [...] authentic”.<sup>316</sup> I have sufficiently explained what ‘authentic reflection’ means earlier in this chapter, let me now recall how the minimal level of competency is characterised:

...the hypothetical reflection we imagine here must be such that it is not the product of social and psychological conditions that prevent adequate appraisal of oneself. This requires that the person have the general capacity to reflect adequately without constriction, pathology, or manipulation. This capacity includes the ability to assess the various aspects of one’s self and conditions, and the freedom from those factors and conditions that we independently know effectively prevent minimal self-understanding. A person who endorses his decisions while in an uncontrollable rage, or while on heavy doses of hallucinogenic drugs, or from having been denied minimal education and exposure to alternatives, does not adequately reflect in this way.<sup>317</sup>

Reflection needs therefore to happen in certain *social* and *psychological* conditions (i.e. it needs to occur in conditions of ‘procedural independence’ – as I explained in previous chapters). Additionally, it is important to notice that competent reflection is understood a means to the end of “adequate appraisal of oneself”, which includes being able to assess “various aspects of one’s self and conditions”.

---

<sup>316</sup> Christman, *The politics of persons*, 146.

<sup>317</sup> *Ibid.*, 146-7.

Regarding the aim of securing an “adequate” evaluation of oneself, I take “adequate” here to imply that one does not need complete self-knowledge (the type only a fully self-transparent subject would have). The level of self-knowledge that is sufficient for the purposes of autonomy falls short of perfect transparency, for Christman. Being aware of one’s desires, preferences, value commitments, are the kind of relevant information one needs to consider when judging the autonomy of one’s dispositions.

It does not seem so clear, however, what could be considered an “adequate” capacity to assess one’s conditions. Some cases immediately strike as exhibiting an “inadequate” capacity, like radical spatial or temporal disorientation (not being able to say at all where one is or what year it is, for example). However, whether one should acknowledge the lack of subtler interpretative capacities to assess “one’s conditions” as an incompetence is not clear. Does one have the capacity to assess one’s conditions, for example, if one does not understand how financial markets work? Or if one is impulsive? Or acts intuitively?

Of course, as I argued earlier, what counts as “adequate” assessment of one’s conditions will be dependent on the potential illegitimate interferences that that one is trying to limit by reflective means. In Christman’s case, as I previously argued, the main heteronomous interferences considered seem disruptive in nature (i.e. they interrupt an agent’s “normal” development). How does this view of illegitimate heteronomy affect his definitions of “adequate” self-assessment and of the set of capacities needed to perform such an assessment?

As I explained earlier, being able to identify sustained features and give coherence to these features is a crucial capacity for Christman. This makes sense in light of an underlying ideal of authenticity. However, such a characterisation of “adequate” assessment becomes ineffective if one challenges Christman’s underlying commitment to an authentic self. If one’s development and one’s sustained commitments could also be part of the problem, then more “disruptive” reflective capacities should also come into the picture.

And what would a set of capacities less oriented towards sustaining stable or settled ways of being include? I do not intend to provide an exhaustive list of all the capacities involved in autonomous reflection or a strict definition of these capacities. Indeed, I believe that doing the latter would be counterproductive for my project. As I argued in Chapter 2, defining competence conditions too strictly, runs the risk of indirectly restricting agents' opportunities for challenging dominant ways of thinking and being. In this sense, my critique of the competence conditions embedded in procedural accounts is mostly negative. Still, it should be clear from my discussion in Chapter 2, and from my proposals, that I do value reflective and critical capacities and that I do consider some minimal abilities necessary to carry out a critical self-assessment.

I would like to suggest that a model of autonomy that is also wary of settled identities and sustained values should at least include two sets of abilities:

First, some abilities are necessary to unmask the contingency of one's sustained ways of being or character traits. Understanding the latter requires certain cognitive abilities (e.g. those necessary to minimally grasp the historical or social function of certain social roles) but also abilities and conditions of other kinds. For instance, not having one's self-worth or idea of one's value exclusively connected to the trait that one is trying to change is key. As the cases that I provide in Chapters 4 and 5 show, if a certain way of being is the *only* way in which I see myself as having (for instance) dignity, it is understandable that challenging this way of being might be resisted. Self-trust also seems crucial to make the task of challenging oneself (i.e. one's stable commitments and character traits) bearable – very vulnerable identities might find it impossible to engage in forms of self-critique (or even self-parody) in a way that does not completely undermine agency (e.g. by making an agent lose any motivation to act at all).

Second, imaginative and creative skills are also key when it comes to defining alternative ways of being. One of the implications of challenging authenticity is that

autonomy becomes much more connected to creative and self-transformative practices than to a form of self-hermeneutics. As Meyers notes, imagination skills “enable individuals to envisage feasible options—to audition a range of self-conceptions they might aspire to and to preview a variety of courses of action they might follow”.<sup>318</sup> Indeed, as the examples provided in Chapters 4 and 5 show, in order to challenge one’s stable commitments, it is necessary to be able to carry out exercises like putting oneself “in the shoes of others” (e.g. to try to adopt the perspectives of others who occupy social roles different to mine), to engage creatively with existing norms and to identify the “room for play” that they might allow for, and even to imagine different ways of being and living.

All in all, my point is that historical models of autonomy as they stand “trust” the information one can get from one’s biography or personal development too much and that, as a result, they value too highly reflective skills like coherence and self-control which are adequate given their commitment with authenticity. As it stands the condition of reflective competence seems to privilege conservative over disruptive outcomes and this is especially worrisome if one wishes to include in this condition the necessary reflective resources to limit subjection.

---

<sup>318</sup> Meyers, *The Feminist Debate over Values in Autonomy Theory*, 121. Christman also seems to suggest that imagination is an element of reflective competence in his analysis of Kaew (Christman, “Coping or Oppression,” 223) but Christman has not, to my knowledge, developed an alternative definition of ‘reflective competence’ to that offered in *The Politics of Persons*.



## Chapter 4: Substantive Accounts of Personal Autonomy

At this point, someone could rightfully wonder if it is sensible to hold on to a proceduralist project. After having discussed the different ways in which procedural accounts of autonomy could be deemed problematic or insufficient, it seems only reasonable to consider alternative approaches to autonomy. This is what I do in this chapter. Specifically, I focus on *substantive* accounts of autonomy and, more specifically, on the feminist objections against procedural accounts that have prompted some contributors to adopt such accounts.

Turning to substantive accounts and to the contribution of feminist philosophers seems all the more sensible if we briefly review the three areas of tension which I identified within proceduralism in Chapter 2. To recall, I argued that: (i) Proceduralist theorists need to define ‘procedural independence’ more clearly and, crucially, they need to avoid a merely “diagnostic” notion of procedural independence – i.e. in addition to clearly stating what circumstances constitute violations of procedural independence, it would be helpful to say more about the way in which agents who are not in optimal reflective conditions could *enhance* their “independence of mind”.

In this chapter, I connect point (i) with the feminist worry that structural oppression can limit one’s ‘psychological freedom’ in a way that remains undetected by procedural models. For example, a key concern for feminist theorists of autonomy has been to determine whether women living in a patriarchal context could have the ‘independence of mind’ that would be required to limit the effects of sexist norms on themselves (e.g. their preferences, desires, and character) and their actions.

Furthermore, I claimed that (ii) historical procedural accounts need to state less ambiguously what kinds of historical materials (e.g. biographical, social, political) need to inform reflection given their aims. Indeed, I have claimed that a narrow understanding of one’s ‘psychological history’ (e.g. one that focuses merely on biographical facts) would leave

agents ill-equipped to face very common potentially illegitimate interferences (e.g. gender oppression).

In relation to this second point, in this chapter I consider the way in which feminist theorists of autonomy have problematized the status of character traits, preferences, and even emotional tendencies that are typically considered “personal”. I focus mostly on the feminist worry that *social* stereotypes and norms affect agents’ psychologies both directly (e.g. by affecting agents’ self-conceptions, preferences, and desires) and, more indirectly, by limiting the life paths available to individuals of a certain “kind” in a particular context.

Finally, I have argued that *(iii)* procedural theorists need to revise their underlying conceptions of oppression. I have claimed that, because proceduralists typically see manipulation and deception as the two paradigmatic forms of illegitimate interference, their proposals are not entirely suitable to respond to subjection insofar as subjection does not require the existence of an identifiable deceiver or manipulator.

Feminist philosophers have made oppression a central worry in their theories and, crucially, they have typically understood oppression as something non-reducible to dyadic one-off interventions. That is, while some forms of patriarchal violence may have an identifiable perpetrator (e.g. sexual assault or psychological violence), one can also face oppressive circumstances where finding an identifiable oppressor or a manipulator is not possible or is even beside the point (because the oppression is structural rather than interpersonal).

Picturing oppression differently has consequences for available procedural models of autonomy. Since procedural independence would be typically ascribed when competent agents have not been exposed to (past or present) illegitimate interventions (e.g. manipulation or coercion), and since structural oppression (as I have already argued) does not necessarily involve any disruptive intervention, then it appears that many agents who are structurally oppressed would *satisfy* procedural independence conditions. Once procedural

independence is guaranteed (as the discussion in Chapters 2 and 3 explained) an agent's autonomy would be typically decided through an 'authenticity' test. Roughly speaking, the latter means that autonomy is granted when self-assessment does not trigger experiences of alienation. In a word, then, when an agent is deemed procedurally independent, then autonomy is decided on the base of a test in which agents' psychologies are "trusted".

The problem with the procedural solution described above, it could be argued, is that it seems to overlook the specifically psychological effects of oppression, especially when oppression can be likened to subjection. As I suggested in Chapters 2 and 3, we should avoid picturing oppression as something which "merely" *interferes with* or *alters* psychological traits like one's sustained preferences and characters – oppression can also shape these sustained preferences and characters from the very beginning. The question then arises of whether trusting the psychologies of those oppressed is the best strategy to push back against oppression.

In order to develop these issues, I take the case of gender oppression as a starting point. It is worth mentioning, however, that taking the case of gender as a starting point does not mean that the conclusions drawn in this chapter only apply to the particular case of gender oppression. As it becomes clearer below, my aim is also to provide tools that are relevant to other forms of subjection, like being socialized in a racist environment or in one where social norms and stereotypes reinforce class oppression. I consider the issue of race oppression through some of the examples in this chapter and the case of class oppression through a more elaborated example developed in Chapter 5.

Many of the substantive positions developed below are then motivated by the same concern: agents' *psychologies* are much more affected and conditioned by oppressive *social* elements than standard procedural models can account for. What makes these positions 'substantive'? Their point is that oppression and its psychological effects cannot be adequately assessed and limited unless one adopts values that are *external* to the psychologies

of the agents concerned. In other words, substantive accounts ask for more than the typical procedural package of internal psychological coherence (e.g. coherence between agents' preferences and their sustained values) plus procedural independence. Crucially, substantive models of autonomy adopt normative commitments that ensure that agents' psychologies develop in the *right* way or in the *right* (normative or relational) conditions. The key idea behind substantive models is that some sort of "robust" normative commitment is necessary to secure autonomy, whereby this commitment may not be one accepted by the agent in question.

Let me briefly introduce the three main ways in which substantive theorists make this move:

First, one could suggest that autonomy should be decided in light of the *actual* socio-relational conditions in which agents live. According to Oshana's view, for example, agents should not be considered autonomous unless they have *actual* control over relevant aspects of their lives, i.e. agents need to be *substantively* independent.<sup>319</sup>

Oshana sets limits to the modes of living that are compatible with autonomy and does so *irrespective* of the psychological stances that individuals may have towards these modes of living (and the procedural independence they may have available). The move in Oshana's model is to give agents' psychologies (e.g. her preferences or affective responses) a *non-decisive* role when determining their autonomy because some ways of living *are* heteronomous *even* when reflectively endorsed by agents under conditions of procedural independence. Indeed, as I show below, the point that oppression may affect one's psychological states is secondary when it comes to identifying heteronomy in Oshana's theory. Consider for example a "deferential wife": Oshana's position is that an agent who lives her life as a deferential wife

---

<sup>319</sup> Oshana, *Personal Autonomy in Society*, 86.

would be heteronomous, not because she might come to *want* to be subservient, but because she *is* subservient.<sup>320</sup>

Second, another possible approach is to keep the focus on agents' psychologies – like proceduralists do – but with the purpose of unveiling inherently psychological aspects of oppression that purportedly slip through the proceduralist net. For Stoljar, oppressive norms and pernicious “symbolic” circumstances (e.g. living in a context where only damaging stereotypes are available for self-definition) could limit one's autonomy even without strictly limiting independent action or even (as Stoljar argues more recently) without affecting one's capacity to (correctly) perceive those norms and stereotypes as false or inadequate. What is more, Stoljar notes that the internalisation of oppressive aspects of one's context may *differently* affect agents who live in *similar* socio-relational conditions. This means that we would be overlooking a relevant and distinctive form of oppression if we focused exclusively on the socio-relational aspects of agents' realities.

A third option is to limit the range of preferences or lifestyles that could be considered autonomous but to do so indirectly, i.e. not by setting restrictions on preferences or lifestyles themselves, but on *self-relations*. The idea behind proposals like this one is that living certain lives or endorsing certain preferences (e.g. those which involve some form of subservience) could be the result of an agent not seeing herself as someone valuable or respect-worthy enough. Moreover, acceptance of problematic forms of dependency could even indicate a failure to see oneself as the ultimate authority vis-à-vis first-person assessments, like it happens when agents underestimate feelings of discomfort or anger because they deem themselves “crazy” or “neurotic”.

For example, Benson argues that gender oppression involves not merely the internalisation of oppressive preferences (e.g. those which derive from oppressive social

---

<sup>320</sup> Ibid., 59.

norms) but it also makes women internalise problematic “attitudes towards themselves”.<sup>321</sup> When oppression has the latter effect, then it is unrealistic to expect that oppressed agents critically assess themselves as procedural models demand them to do. Benson’s point is that a distorted sense of one’s own value will limit the critical potential of a procedural test.

Moreover, Benson argues that historical procedural accounts are not necessarily better positioned to tackle the internalisation of problematic attitudes towards oneself. For example, the information on the origin of one’s values and preferences might not make critical self-assessment easier for a woman living in a patriarchal context, Benson claims:

A woman could be so unsure about her own value apart from men’s attitudes toward her that she would be content to know that, for the most part, the pressure of male interests led her to be this way.<sup>322</sup>

So, according to this third strategy, it is essential that we make sure that agents self-relate in the *right* kind of way for procedural tests (including historical ones) to be trustworthy.

Going back to the commonalities of substantive accounts, the three abovementioned strategies share (in spite of their differences) a crucial feature: they are willing to restrict what it means to be autonomous or to live autonomously in a way that may be *explicitly* content-laden. For example, referring back to Meyers’ distinction between a ‘Directivity Axis’ and a ‘Constitutivity Axis’ (presented in Chapter 1), we can say that substantive accounts typically import normative content into their models directly through the Directivity Axis, i.e. they “prescribe or proscribe certain types of behaviour”.<sup>323</sup> This means that certain lifestyles,

---

<sup>321</sup> Paul Benson, “Autonomy and Oppressive Socialization,” *Social Theory and Practice* 17, no. 3 (Fall 1991): 391.

<sup>322</sup> *Ibid.*, 394.

<sup>323</sup> Meyers, “The Feminist Debate Over Values in Autonomy Theory,” 115. This claim applies in particular to Stoljar’s and Oshana’s theories, as it will become clearer below. Meyers considers Benson’s account (which in the literature is typically labelled a “weakly substantive” one) as neither ‘value-laden’ nor ‘value-saturated’ on the Directivity Axis but as “value-utilizing” on the Constitutivity Axis. (*Ibid.*, 120) In this view, Benson’s account would not be that different from many procedural accounts. I share the motivation behind this analysis, as it emerges from Chapters 2 and 3. However, it still remains true that the three (substantive) accounts which I analyse in this chapter are different from procedural accounts insofar as they *explicitly* rule out certain values, attitudes or ways of living as incompatible with autonomy. This justifies analysing these three accounts together.

social norms, social conditions, social relations or forms of relating to oneself are deemed *objectively* heteronomous or conducive to heteronomy. In a word, for the three authors introduced above (and others adopting similar positions), autonomy is only possible under specific socio-relational, normative or psychological conditions that apply *regardless* of whether reflectively competent agents might prefer, desire or judge otherwise. My aim in this chapter is to elucidate the scope and force of this substantive point and, ultimately, to shape my own proposal in light of the valuable lessons that can be obtained from the different strategies used to argue in favour of this claim.

This chapter is structured as follows:

In the first section, I focus on the second strategy mentioned above, namely I consider the charge that available procedural models as they stand are insensitive to ‘psychological oppression’. That is, I am interested in the arguments put forward to support the claim that agents subjected to, for example, oppressive gender norms seem to “counter-intuitively” qualify as procedurally independent. Claiming the latter – I argue – does not force us to automatically reject the procedural project altogether. Indeed, signalling that available procedural tools are not sensitive to common forms of oppression is a serious pitfall but not necessarily an unsurmountable one. However, I show that the substantive theorists that I analyse below *do* adhere to this stronger thesis – namely, they argue that there is something *particularly* insidious about harmful stereotypes and oppressive social norms that forces us to adopt a *different* (i.e. non-procedural) project to secure autonomy. To unpack and justify this claim, I consider two different hypotheses held by theorists who adopt substantive accounts – both focus on the psychological harms done by oppression but they do so differently:

The first hypothesis is that living in oppressive normative conditions necessarily “clouds” one’s normative capacities – i.e. one’s ability to distinguish good from bad

norms.<sup>324</sup> Because one has internalised an oppressive normative context – the argument goes – one loses (in ways that become clearer later) the capacity to reflectively resist or challenge pernicious norms. Consequently, the preferences and desires that derive from those norms should not be taken as being necessarily an expression of the agents’ wills (or of their authentic selves). This first version of the substantive critique can be summarized as follows:

*S<sub>i</sub>*: A substantive account of autonomy is the only viable approach under conditions of (what I call) ‘subjection’ because, in contexts of subjection, defining autonomy as ‘authenticity + procedural independence’ would trivially – and wrongly – deem most agents autonomous. Why? Because internalising the norms and stereotypes available in oppressive contexts would make these stereotypes and norms impossible or almost impossible to reject, resist or assess. If endorsement or non-alienation is the most likely outcome in these conditions, then the procedural solution does not meaningfully distinguish between autonomous and heteronomous agents.

I show that this thesis, which underlies the work of some substantive theorists, does not do justice to the study cases that are typically put forward to sustain it (e.g. Luker’s cases, analysed below). Moreover, as I argue, the notion of ‘internalisation’ does not fully capture oppression as ‘subjection’. As I explained in Chapter 2, subjection to social norms or “truths” happens both by way of ‘technologies of power’ (which continuously aim at normalizing and regularizing individuals and groups of individuals) and of ‘technologies of the self’ (through which individuals make themselves subjects of particular codes). In a word, subjection to problematic social injunctions or moral codes does not happen as a one-off event that clouds

---

<sup>324</sup> This “normative capacity” should not be mistaken for “reflective competence”. As I explain below, the substantive point is that agents might be competent (in the sense of being, e.g., minimally rational) and still lack the capacity to assess the truth-value or relevance of the norms that inform their reflection.



one's judgement forever and this, in spite of making matters more complex, also multiplies the opportunities for limiting subjection.

In the second section, I consider a second version of the substantive critique that overcomes the objections raised against  $S_1$  and that is still strong enough to challenge procedural accounts. The second version of the substantive thesis could be put as follows:

$S_2$ : A substantive account of autonomy is the only viable approach under (what I call) 'subjection' because, *even if* an agent could take a critical stance on the social norms and stereotypes that influence her preferences (and on these preferences, themselves), these problematic norms and stereotypes will *still* constrain one's psychological life in many illegitimate ways. In a word: oppressive norms or stereotypes will limit one's psychological freedom, and therefore, one's autonomy, *regardless* of one's stance on them.

Substantive theorists who worry about the effects of psychological oppression could argue for the need to secure *substantive* independence by committing to  $S_2$  alone. However, I argue, even if we acknowledge that someone affected by oppressive stereotypes is less autonomous than someone who is not (as Stoljar rightfully claims), it seems that we judge too quickly if we say that being able to reflectively problematize and resist these stereotypes does not make agents less heteronomous and that, therefore, this constitutes no meaningful form of self-government.

My point is therefore that, when what we understand by 'oppression' can be likened to 'subjection', then the ability to see a preference or character trait for what it is (e.g. a motivation derived from a potentially pernicious social injunction) *is* pushing back on heteronomy. Crucially, since this move could lead, as I have argued before, to challenge features that one would have otherwise deemed "authentic", then it seems that critical assessment informed by the *social* nature of one's character could contribute to a more robust procedural self-assessment.

Admittedly, the suggestion just made will not do enough to *fully* address substantive theorists' worries, namely that we cannot fully tackle oppression through a *reflective* procedure alone. I am aware of this fact and, as I mentioned before, I do not find the strategy of defining oppression too narrowly, naively, or artificially, helpful at all. Indeed, adopting a one-sided view of oppression risks falling back on the oversimplified understanding of oppression as 'internalisation' of an oppressive context that I problematized above.

Instead, my proposal is that substantive theorists are right when they claim that "procedural independence" in *certain* oppressive contexts would fail to secure autonomy when the latter is seriously threatened by external (normative or material) constraints. However, my point is also that oppressive contexts which are mainly characterised by oppressive norms and stereotypes (like the ones that Stoljar is particularly worried about and, to an extent, Benson) are *precisely* those contexts in which the appropriate independence of mind *could* make a *qualitative* difference vis-à-vis limiting one's oppression. More importantly, I propose that this qualitative difference that reflectively challenging stereotypes could make, could progressively lead to the more external, social, and relational change that substantive theorists (rightfully) perceive as necessary to fully tackle oppression.

I accept, nonetheless, the point that in *some* oppressive circumstances no 'independence of mind' could be enough to secure autonomy insofar as it will not guarantee the necessary control over one's life and the access to significant options that is crucial for meaningful self-government. I set this demarcation after looking into Oshana's theory and I argue that a minimal 'substantive independence' condition, i.e. 'non-domination', is necessary to exclude those cases in which oppression cannot be significantly limited through procedural means. Therefore, I propose that an exclusively external *socio-relational* account of autonomy (i.e. one in which reflection does not play a decisive role) is necessary only under

cases of ‘domination’ – understood in the sense defined in Chapter 2. However, a revised procedural strategy is still adequate to tackle forms of ‘subjection’.

My aim in this chapter is then to disentangle the different ways in which individuals may be affected by structural oppressive conditions and to decide on the strategies that are most suitable to deal with different oppressive conditions. I start making the case that a two-tracked model of autonomy is necessary to limit oppression: as I anticipated, the model I develop in Chapter 5 is procedural for contexts of subjection and substantive for contexts of domination.

### I. The Problem of “Fully-Internalised” Oppression

One way to problematize procedural models consists in showing that oppression has psychological effects that could not be detected by the first-person assessment of the oppressed. For the sake of clarity, I illustrate this worry through Stoljar’s (famous) ‘feminist intuition’, which she formulates based on an analysis of Kristin Luker’s study of women who take contraceptive risks.

Before beginning a discussion on the ‘feminist intuition’, it is worth mentioning that Stoljar does not consider the procedural project in itself incompatible with a feminist agenda. Crucially, Stoljar notes that proceduralism *is* compatible with a *relational* picture of autonomy, which is especially relevant for feminists.<sup>325</sup> Therefore, the debate between defendants of proceduralism and defendants of substantive accounts should not be framed as one opposing “self-sufficiency” to relational views on agency and autonomy.

Indeed, as I have explained in previous chapters, it would be inaccurate to assume that adopting a procedural view amounts to assuming that, for example, “self-sufficient” agents

---

<sup>325</sup> Stoljar, “Autonomy and the Feminist Intuition,” 95.

(if they exist) would be more autonomous than those who are not. In fact, personal relations can also have a key role in procedural accounts: proceduralism does *not* require that an agent becomes a reflective agent on her own or, in other words, proceduralism is compatible with a *causally* relational picture of autonomy.<sup>326</sup>

Furthermore, according to Stoljar, procedural ‘content-neutrality’ can be “congenial” to feminist purposes.<sup>327</sup> In procedural terms, it is technically possible to autonomously value and maintain attachment to others and, even, dependency on others.<sup>328</sup> For example, as we saw in Chapter 1, procedural models could theoretically account for the autonomy of preferences which entail or maintain dependency on others (as it happens, say, when one’s wellbeing or fulfilment is tied to someone else’s wellbeing or fulfilment) as long as the concerned agent follows the relevant reflective-validation procedures.

Similarly, when Stoljar and other substantive theorists argue that some form of ‘substantive independence’ (e.g. psychological or external) is necessary for autonomy, what is implied is *not* that autonomous agents should be “radically” independent from others. Indeed, Stoljar claims that “[t]he question for *all* theories of autonomy is *what kinds* of socialization are incompatible with autonomy”.<sup>329</sup> This, as I have just reported regarding procedural theories but want to extend to substantive ones, does not exclude that *some* kinds of socialization are not only compatible but also *necessary* for autonomy.

Stoljar’s key worry could be then summarized as follows: even if proceduralism is *theoretically* compatible with feminist concerns and with a relational picture of autonomy (i.e. with one in which others play a key role in one’s autonomy), it is, nonetheless, *practically* ill-equipped to detect and limit crucial social influences that should be deemed illegitimate from

---

<sup>326</sup> I take the term ‘causally relational’ from Westlund, “Rethinking Relational Autonomy,” 27. See Christman, *The Politics of Persons*, 9 as an example of this.

<sup>327</sup> Stoljar, “Autonomy and the Feminist Intuition,” 95.

<sup>328</sup> *Ibid.*

<sup>329</sup> *Ibid.*, 97; emphases added.

a feminist point of view. According to Stoljar, social influences which are incompatible with feminism *should* make agents count as *heteronomous* and this is what she calls the ‘feminist intuition’, which is “precisely the intuition that the subjects are not autonomous”.<sup>330</sup> Therefore, a failure to deem agents affected by oppressive gender norms heteronomous, shows the insufficiency of procedural accounts.

Stoljar characterises the feminist intuition further as follows:

In certain cases, even preferences satisfying the standards of critical reflection that are required by procedural accounts would still be regarded as nonautonomous by many feminists. This is because such preferences are influenced by pernicious aspects of the oppressive context. They therefore attract what I call *the feminist intuition*, which claims that preferences influenced by oppressive norms of femininity cannot be autonomous.<sup>331</sup>

This passage is not completely clear. To be sure, the problematic point for Stoljar is not that agents are influenced by their contexts since, as I have argued above, Stoljar would accept that *all* agents are influenced by their contexts. Therefore, the issue is to clarify whether Stoljar considers autonomy-undermining to be influenced by *oppressive* contexts *tout court* (i.e. in a way that would equally apply to, say, feminists or other activists) or if the threat is the possibility of judging and acting while under the influence of the “*pernicious* aspects of the oppressive context”.<sup>332</sup> How are these positions different and why does it matter? I explain this as follows:

It is possible to think of positive relations and even valuable “skills” that arise from oppressive contexts. For example, forms of solidarity amongst the oppressed could be seen

---

<sup>330</sup> Ibid., 96.

<sup>331</sup> Ibid., 95; original emphasis.

<sup>332</sup> Ibid; emphasis added.

as the “result” of the oppressive context in which they emerge. Should we deem these forms of solidarity heteronomous? Moreover, as I suggested in Chapter 2, being ‘subjected’ – i.e. being “shaped” as a subject by a context and its moral codes – does not merely mean that one has been *passively* made a certain agent, but also that one has acquired certain agentic skills to *actively* constitute oneself as a subject of a certain moral code.<sup>333</sup> Could these skills contribute to enhance one’s autonomy even when they “derive” from oppressive contexts?

For example, oppressive gender norms that encourage women to conform to specific aesthetic and moral standards (i.e. that encourage them to become a certain kind of person) presuppose that women are *agents* capable of a number of actions in the world and on themselves. After all, even to become an agent who conforms to available gender stereotypes, it is necessary to engage in activities like self-monitoring, self-transformation, and self-control which require a degree of “skilfulness”.<sup>334</sup> These capacities or skills may not be *inherently* pernicious or oppressive – indeed, one could engage in forms of self-monitoring and self-transformation with the purpose of “being otherwise” (i.e. being different from what is socially expected of oneself in an oppressive context). Additionally, one could even engage with available norms or stereotypes in a way that does not merely reinforce one’s oppression. I provide examples of this below, but for the sake of clarity let me unpack and clarify Stoljar’s position first.

---

<sup>333</sup> See Chapter 2, section III on page 102.

<sup>334</sup> See for example Bartky’s (critical) analysis of the disciplinary technologies involved in the production of feminine bodies. Bartky does not defend the thesis that the skills acquired as a result of discipline could enhance women’s autonomy but, nonetheless, she spots a tension between the aim of empowering women and the critique of forms of agency enabled by discipline. Bartky claims: “While its imposition may promote a larger disempowerment, discipline may bring with it a certain development of a person’s powers. Women, then, like other skilled individuals, have a stake in the perpetuation of their skills, whatever it may have cost to acquire them and quite apart from the question whether, as a gender, they would have been better off had they never had to acquire them in the first place. Hence, feminism, especially a genuinely radical feminism that questions the patriarchal construction of the female body, threatens women with a certain de-skilling, something people normally resist: Beyond this, it calls into question that aspect of personal identity which is tied to the development of a sense of competence.” (Bartky, “Foucault, Femininity, and The Modernization of Patriarchal Power,” 77.)

In what follows, I argue that Stoljar seems to assume the strongest alternative mentioned earlier, namely Stoljar considers that even acts of resistance within a pernicious normative environment should be deemed heteronomous. As I show below, the possibility of acting or feeling in a way that counters an oppressive normative environment is given more place in Stoljar's recent work but this possibility continues to be, nonetheless, insufficient to grant autonomy.

To understand Stoljar's worries it is useful to take a closer look at her analysis of Luker's study. This case study was carried out in California during the early 1970's and collects interviews with women who face unwanted pregnancies and who decide to undergo an abortion.

The study is interesting from a feminist perspective for many reasons:

First, Luker's analysis shows how gender norms and stereotypes influence not only women's decision-making and bargaining processes but also the way in which medical professionals and researchers tackle issues of contraception and abortion. Crucially, her point is that "much of our [i.e. social researchers'] thinking on sex, abortion, and contraception has been based not on fact but on value judgements."<sup>335</sup> Having said this, however, it is important to note that Luker does not attribute any subjective intentionality to the "bias" in sexual-health research or medical practice. That is, Luker's point is that all social actors (i.e. the subjects of her study, their partners, and the medical and research communities involved) are somehow "caught" in a normative framework that is pernicious to women and that thus needs to be questioned.

Luker's project is, therefore, 'political' and intends to "analyse why people of goodwill, working in the presumed best interests of all of us, have helped create social

---

<sup>335</sup> Kristen Luker, *Taking Chances: Abortion and the Decision Not to Contracept* (Berkeley: University of California Press, 1975), ix.

realities which are oppressive to women.”<sup>336</sup> The latter is highly relevant because Luker’s analysis should be regarded as a structural one, and not as an exclusively relational or interpersonal one. In other words, Luker’s study does not typically depict instances of *dyadic* interference (e.g. of illegitimate manipulation or deception) but, rather, of ‘subjection’. Indeed, the study describes the actions, strategies, and ambivalences of individuals who are subjected to social injunctions to become and to live as certain kinds of subjects (i.e. as ‘women’) and it records some of the medical, familial, and social mechanisms involved in this process.

Second, this study is valuable from a feminist perspective because the profile of Luker’s subjects makes it difficult to attribute their “bad” contraceptive choices to their “immaturity” (subjects were not particularly young) or to their reflective “incompetence” (subjects participating in the study were minimally educated).<sup>337</sup> Therefore, Luker’s study allows to problematize constraints to agency that are not due to *personal* (reflective or emotional) factors but, rather, to the *social* normative circumstances in which women make decisions about pregnancy, about their sexualities, and about their lives more generally. (Indeed, one of Luker’s key moves is to dispose of the overly reductive idea according to which individual women’s contraceptive behaviours are *exclusively* aimed at avoiding pregnancy; rather she unveils the many social injunctions and stereotypes that play a role when women make contraceptive decisions).<sup>338</sup>

Crucially, Luker’s study departs from other views on ineffective contraceptive usage and abortion that (rightfully) strike feminists as problematic. According to Luker, the tendency before her study had been to explain ‘contraceptive risk-taking’ by focusing exclusively on women’s personal and intellectual “qualities” or lack thereof. Previously,

---

<sup>336</sup> Ibid.

<sup>337</sup> Ibid., viii.

<sup>338</sup> Ibid., 16.



women who took contraceptive risks were deemed “at best irrational and at worst pathological.”<sup>339</sup> For example, one of such previous views claimed:

If it is possible ... to characterize the personality that will fail most consistently in contraceptive usage, it is the immature, dependent, self-punishing individual who has a feeling of low self-esteem and self-worth and has little, if any, desire to control his or her life.<sup>340</sup>

Contrary to this previous tendency, Luker does not present the subjects of her study as “deficient” agents but, rather, she highlights the fact that they were “both reasonable and logical given their *own* definition of the situation”.<sup>341</sup>

The latter brings in an important point to the discussion, namely how formal or substantive we want our definitions of what is reasonable to be. In the latter case, one might be ready to deem some “risky” attitudes as unreasonable on the grounds that, though internally coherent, they do not seem to be in line with “objective standards of ‘Reason’”.<sup>342</sup> Take the case of smokers: someone’s decision to smoke might be internally coherent if it shows, say, a consistent assessment of preferences and desires. However, this decision might strike some as being grounded on *bad* reasons, namely ones that reasonable people *should* reject. However, in order to make the latter claim, one needs to resort to ideals of, for instance, what is and what is not *objectively* in people’s best interest, which can and often diverges from what people actually pursue. Therefore, substantive views of rationality may require us to commit to ideals of the reasonable and the good that transcend those that agents may have set *themselves*.

---

<sup>339</sup> Ibid., 16

<sup>340</sup> Sandberg and Jacobs in Kristen Luker, “Contraceptive Risk Taking and Abortion: Results and Implications of a San Francisco Bay Area Study,” *Studies in Family Planning* 8, no. 8 (Aug., 1977): 191.

<sup>341</sup> Luker, *Taking Chances*, 17; emphasis in original.

<sup>342</sup> Susan Wolf in Stoljar, “Autonomy and the Feminist Intuition,” 97.

As Stoljar notes, the debate between formal vs. substantive accounts of rationality finds “structural parallels” with the debate between procedural and substantive accounts of autonomy.<sup>343</sup> In this sense, we could say that the feminist intuition (i.e. that Luker’s agents *are* heteronomous)<sup>344</sup> is triggered by normative commitments that are *external* to the ones that the agents in the study hold. Moreover, Luker’s study shows the insufficiency of procedural accounts, Stoljar would argue, because these accounts cannot account for the *fact* that these agents are non-autonomous *in spite of* being competent (i.e. minimally rational and coherent) and satisfying other procedural conditions.

Take competence conditions, for example. Stoljar mentions that Luker’s subjects decided not to use contraception following a consistent bargaining process in which the costs and benefits of contraception were well assessed. Indeed, Stoljar highlights that the subjects were actually quite thorough in their cost-benefit analysis: they considered the medical, biological, economic, and social costs and benefits of contraception.<sup>345</sup> Therefore, Luker’s subjects would qualify as minimally rational according to content-neutral procedural accounts.<sup>346</sup>

Second, in addition to being competent, Stoljar claims, the agents in Luker’s study would satisfy other conditions set by procedural models of autonomy, such as reflecting in a way that is free from inhibiting factors or involves endorsing their preferences. I do not detail here the different variations of procedural conditions that Stoljar considers.<sup>347</sup> For the sake of the argument, I concede that many of Luker’s agents seem to satisfy the different conditions that different procedural accounts set for autonomy and, crucially, I accept that

---

<sup>343</sup> Ibid., 97.

<sup>344</sup> Ibid., 96.

<sup>345</sup> Ibid., pp. 98-9.

<sup>346</sup> Ibid., 97.

<sup>347</sup> See Ibid., pp. 100-7. She analyses five different conditions extracted from procedural accounts and shows that the subjects in Luker’s study would satisfy all of them – that is, they would successfully: engage in counterfactual reflection, show coherence in their self-assessment, endorse the contents under scrutiny, show self-knowledge, and reflect in a way that is free of inhibiting factors.

these agents would satisfy the conditions set by the historical models that I have analysed so far.

The reason why Luker's agents seem to satisfy the conditions set by historical procedural accounts is that (at least at the time when they make the decision not to use contraception) these agents appear not only competent but, furthermore, they seem to act in an *authentic* manner. For example, Luker's agents do not report experiencing, at the time of the events, 'acute alienation' (in Christman's sense) vis-à-vis their preferences and decisions (even if, as I explain later, they do manifest some "conflict"). That is, Luker's subjects generally "endorsed" their preferences and understood them (indeed, some of them even explicitly claim that it was all "worth it").<sup>348</sup> The issue is rather that, once pregnancy becomes a reality, some of the participants in the study felt that their past expectations were not met or that they needed to deal with unforeseen complications.<sup>349</sup> All in all, my point is that claiming that Luker's agents acted inauthentically seems inaccurate.

The intuition that Luker's agents are non-autonomous arises, according to Stoljar, when we consider their analysis of the cultural and social pros and cons of contraception. For example, Stoljar notes that "public censure" was one of the costs of contraception considered by some of Luker's subjects.<sup>350</sup> That is, deciding to use a contraceptive method could be read as "planning for sex" and this, given some of these woman's situations in their social environments (e.g. being single), could make them the targets of others' negative judgements.<sup>351</sup> Furthermore, Stoljar describes how some of the participants of the study saw a potential pregnancy as something that, given their contexts, could increase their "value" by, for instance, proving their fertility. Similarly, pregnancy was reported as something

---

<sup>348</sup> Luker, *Taking Chances*, 51.

<sup>349</sup> *Ibid.*, 71.

<sup>350</sup> Stoljar, "Autonomy and the Feminist Intuition," 98-9.

<sup>351</sup> See Luker, *Taking Chances*, 46-51 for examples.

positive on the grounds that it may force an “uncommitted” partner to enter a stable relationship.<sup>352</sup>

The issue, Stoljar argues, is that the costs and benefits described above presuppose *false* social norms. According to Stoljar, these women implicitly accept social dictums like “It is wrong to engage in premarital sex” or “Pregnancy and childbearing promote one’s worthiness”.<sup>353</sup> By “implicitly” I mean that, even when agents might not explicitly endorse these social norms themselves, they do accept preferences that *derive* from these norms.<sup>354</sup>

Stoljar’s analysis does not merely apply to problems associated with women’s agency and socialization. Indeed, Stoljar concedes that the same intuition (i.e. that these agents are heteronomous) could be triggered by other cases in which “formal” reasoning is influenced by problematic norms. Stoljar gives the example of the different reactions we could have towards voluntary smokers. Consider, for instance, that after a well-informed deliberation, an adult decides that the potential negative costs of smoking are compensated by the pleasure that she gets from it.<sup>355</sup> Stoljar argues that, even if one might judge this agent’s decision a “poor” one, our reactions vis-à-vis this happy smoker would differ significantly from those triggered by a teenager who weighs the costs and benefits of smoking influenced by a social norm which says that smoking makes her look, e.g., “glamorous”.<sup>356</sup>

In order to assess more accurately the *scope* and the *force* of the abovementioned substantive critique of procedural models, it is crucial that we consider *how* and to *what extent* pernicious social norms could affect one’s preferences and desires. A key element to consider in light of the latter aim is whether the processes by virtue of which preferences are negatively

---

<sup>352</sup> Stoljar, “Autonomy and the Feminist Intuition,” 98.

<sup>353</sup> *Ibid.*, 99.

<sup>354</sup> *Ibid.*, 109.

<sup>355</sup> *Ibid.*, 99.

<sup>356</sup> *Ibid.*

affected by oppression affect these preferences *irredeemably*, i.e. in a way that could not be reverted by critical reflection. Therefore, it is useful to say more about the connections between oppression and oppressed agents' psychologies.

Stoljar follows Ann Cudd in distinguishing “direct” and “indirect” psychological harms of oppression:

Direct harms are due to the intentional activities of an oppressor group, whereas indirect harms are generated within the psychology of the oppressed themselves: they “occur when the beliefs and values of the privileged or oppressor groups are subconsciously accepted by the subordinate and assimilated into their self-concept or value/belief scheme.”<sup>357</sup>

So, in oppressive circumstances, the psychologies of the oppressed can be damaged both by the actions of others (e.g. humiliation) and by operations that happen *within the oppressed agents' psychologies*. In the latter case, the value frameworks and perspectives of the oppressors are reproduced or enforced by the oppressed themselves without the need of *actual* (further) interventions. Cudd claims that “[t]he paradigms of indirect psychological harms of oppression are shame, false consciousness, and deformed desire.”<sup>358</sup>

In a word, the problem at stake is that of the ‘internalisation’ of oppression, namely the process through which oppressive values and perspectives may become a part of the psychologies of the oppressed. Stoljar explains:

Agents who are oppressed come to internalize their oppression: they come to believe in the ideology of oppression and to make

---

<sup>357</sup> Ann E. Cudd in Natalie Stoljar, “Living Constantly at Tiptoe Stance: Social Scripts, Psychological Freedom, and Autonomy,” in *Personal Autonomy and Social Oppression: Philosophical Perspectives*, ed. Marina Oshana (New York: Routledge, 2015), 116.

<sup>358</sup> *Ibid.*, 117.

choices, and form preferences and desires, in the light of that ideology.<sup>359</sup>

Because of the internalisation of oppressive values and perspectives, the argument goes, oppressed agents may start thinking, feeling, and self-managing in ways that are functional to their oppression.

However, claiming that oppressed agents may internalise oppressive values and perspectives does not conclusively exclude proceduralism as a useful strategy to define and identify autonomy. For example, one may argue that if internalisation affected only some values or if it was reversible (e.g. in light of information on the power-biased acquisition of one's values), then some forms of reflective self-assessment could still have some emancipatory potential like, say, "filtering" values to identify those which are *not* the result of internalised oppression.

The kind of solution suggested above, however, downplays some of the points made by defendants of substantive accounts and, indeed, some of the arguments I presented in Chapter 3 (regarding the difficulties in accessing a "non-oppressed self"). Importantly, substantive theorists highlight that oppression *also* affects the values and some of the capacities that would be necessary to "filter" preferences or values in light of, e.g., their history. For example, as my earlier reference to Benson suggested, internalised oppression could even affect one's self-assessment in light of one's history, insofar as having internalised one's oppression could also have an impact on the personal histories, influences, and forms of socialisation that one will consider acceptable. Therefore – substantive theorists argue – rejecting proceduralism in light of the possibility of internalized oppression is justified because, when one's oppression has been internalised, taking a *critical* stance on internalised values and preferences will be extremely difficult or even impossible.

---

<sup>359</sup> Stoljar, "Autonomy and adaptive preference formation," 227-8.

It seems, however, that there is a small tension between, on the one hand, Stoljar's claim that agents like those involved in Luker's study are rational and competent and, on the other hand, Stoljar's suggestion that there is a form of competency that these agents lack after all. In other words, in which sense are Luker's agents "incompetent"? What is the kind of assessment that they cannot (and could not) achieve?

Stoljar speaks about '*normative competence*', a term she borrows from Benson which refers to an agent's capacity to assess relevant social norms.<sup>360</sup> According to Stoljar, two versions of the '*normative competence*' criterion can be extracted from (different stages of) Benson's work.

First, in Stoljar's reading of Benson, an agent exhibits normative competence in a *weak* sense when agents see themselves as competent enough to justify their actions in light of the normative demands of their context and of others in this context.<sup>361</sup> As Stoljar rightfully notes, this sense of normative competence connects, in Benson's most developed theory, with a sense of one's "worthiness". On these grounds, Stoljar concedes that Luker's subjects do exhibit weak normative competence insofar as they "do not exemplify a level of breakdown of the sense of self-worth sufficient to undermine their capacity to regard themselves as members of the normative community".<sup>362</sup> In other words, the heteronomy of Luker's subjects could not be explained thanks to the third substantive strategy presented in the introduction to this chapter. (To recall, this position defended – amongst others – by

---

<sup>360</sup> Stoljar, "Autonomy and the feminist intuition," 107-8.

<sup>361</sup> Ibid., 107.

<sup>362</sup> Ibid., 109. This assessment seems valid for the majority of the cases reported by Luker. It is worth noting, however, that Benson rejects Stoljar's strategy (i.e. granting Luker's agents 'weak normative competence' in a blanket way) on the grounds that the attitudes of some of the women in Luker's study "manifest marked disengagement or dissociation from their conduct." (Paul Benson, "Feminist Intuitions and the Normative Substance of Autonomy," in Taylor, *Personal autonomy*, 128). This disagreement is not relevant for my argument for now. Indeed, if one were to deny 'weak normative competence' to some of Luker's subjects, then these cases would fall within the realm of Benson's current model of which I say more in the last section of this Chapter.

Benson, attributes autonomy only to agents who see themselves as valuable and as authoritative sources of claims.)

Second, Stoljar extracts an additional normative competence criterion from Benson according to which an agent has *strong* normative competence when she is able to “criticize courses of action competently by relevant normative standards.”<sup>363</sup> According to Stoljar, the internalisation of “false” norms constitutes a clear case of ‘strong normative incompetence’ because:

Women who accept the norm that pregnancy and motherhood increase their worthiness accept something false. And because of the internalization of the norm, they do not have the capacity to perceive it as false.<sup>364</sup>

So, according to Stoljar, agents who (inadvertently or unknowingly) govern themselves in light of *false* norms are normatively incompetent in this second (strong) sense.

Let me note that, even if I do get what (I believe) motivates Stoljar’s claim (i.e. there is something deeply problematic and, indeed, *oppressive* about the norms of femininity that Stoljar refers to), it seems weird to claim that these norms are “false”, especially when this is done in connection with Benson’s (early) normative competence criterion which, as Stoljar correctly notes, refers to “relevant” social norms. It is true that Luker’s subjects (and others) may be better off without these social norms but the latter are not “false” in a crucial sense: these *are* the norms according to which these women are likely to be judged and assessed in their normative communities. In other words: failing to comply with these gender norms would have very real consequences in the social world in which these agents live.

Indeed, if we turn to Benson’s early account, it is not entirely clear if we should deny (strong) normative competence to Luker’s agents. Benson defines the ability to “criticize

---

<sup>363</sup> Benson in *Ibid.*, 107.

<sup>364</sup> *Ibid.*, 109.



courses of action competently by relevant normative standards” (to which Stoljar refers in her analysis of ‘strong normative competence’) as an ability to “appreciate” *relevant* values.<sup>365</sup> Benson provides this definition in the context of a discussion about ‘free action’. According to Benson, actions are free when they are “potential vehicles for self-disclosure” which means that “our free acts are suitable grounds on which certain sorts of evaluative appraisals of us may be formed”.<sup>366</sup> For example, if I do something that could be deemed “kind” or “rude” according to the normative standards applicable to my action in a given social context, then I might be also disclosing that I *am* “kind” or “rude”. Crucially, according to Benson, this disclosure happens *only* when I act in awareness of the way in which my action will be interpreted by others and, therefore, of what my action will disclose about me. In other words, according to Benson, for an action to be ‘free’ it needs to be “one’s own”, and the latter is only possible when one “understand[s] well enough” the evaluative claims that attach to one’s action.<sup>367</sup>

I do not wish to enter a discussion about the details of Benson’s early view, my intention is more modest: I want to question that it is impossible to claim that Luker’s agents are normatively competent in this strong sense too. Indeed, it could be argued that it is their awareness and their understanding of the norms that (they know) *do* apply to their sexual lives in their normative contexts, and will likely motivate judgements about them as persons, which explains their actions.

Admittedly, as I acknowledged above, Stoljar wants to distinguish between *applicable* and *good* normative standards attached to one’s actions. Indeed, Stoljar refers to Benson’s example of an eighteen-year-old woman who –as a result of having internalised social beauty standards– cannot avoid seeing her physical appearance as “deficient”, as a clear case of

---

<sup>365</sup> Paul Benson, “Freedom and Value,” *The Journal of Philosophy* 84, no. 9 (September 1987): 469.

<sup>366</sup> *Ibid.*, 481.

<sup>367</sup> *Ibid.*, 478.

someone who *lacks* strong normative competence.<sup>368</sup> The key issue, Benson and Stoljar would agree, is that –as a result of both her upbringing and the false normative standards of her environment– this woman lacks a capacity to *challenge* the misconception that her (and most women’s) “natural physical appearance is deficient”.<sup>369</sup>

Then, Stoljar’s worry is not so much that Luker’s agents may or may not understand the norms that *actually* apply to them when they act but, rather, that they are not capable of contesting the personal assessments that others (and probably also themselves) will infer from their actions.<sup>370</sup> What Stoljar deems a necessary condition for ‘strong normative competence’ in this case is a capacity to (cognitively and affectively) interrogate and reject the sexist nature of the applicable normative standards. Since Stoljar submits that interrogating an oppressive normative environment is not possible when agents have internalized the norms available in that normative environment, then agents who have internalized sexist or racist social norms will, in Stoljar’s view, typically lack strong normative competence. (My use of ‘strong normative competency’ from now on follows Stoljar’s).

To summarize my points so far, according to Stoljar, Luker’s subjects satisfy all the conditions to be deemed autonomous in a procedural sense *but* they are normatively incompetent in a strong sense insofar as they accept “false” (i.e. oppressive) norms. The latter entails an incapacity to critically assess the oppressive norms that they presuppose and the preferences that derive from these norms. Moreover, since Luker’s agents could not, in the oppressive normative context in which they live, inform their actions by non-oppressive norms, their psychological freedom to form non-oppressive preferences, desires, and self-assessments is restricted. As a result, even if they are capable of self-assessment and reflect

---

<sup>368</sup> Benson in Stoljar, “Autonomy and the feminist intuition,” 108.

<sup>369</sup> Stoljar, “Autonomy and the feminist intuition,” 108.

<sup>370</sup> Admittedly, the issue at stake is not exclusively that of being able to contest what my actions disclose about me in a given normative environment. Additionally, Agents affected by oppressive norms and stereotypes may have a diminished control over their “self-presentations” regardless of whether they reject the norms and stereotypes “relevant” in their social contexts. I come back to this issue in the next section of this chapter.

in a way that is free from directly inhibiting factors like manipulation – i.e. even if they are procedurally independent – they *cannot* limit oppression through subjective means.

From all the above, the substantive thesis  $S_1$  mentioned earlier follows, namely that a substantive account of autonomy is the only viable approach under conditions of ‘subjection’ because, in these contexts, the internalisation of the negative norms and of the damaging stereotypes available would make these stereotypes and norms impossible or almost impossible to reject, resist or critically assess.

In the remainder of this section, however, I want to suggest that holding  $S_1$  appears inadequate or excessive upon closer examination of Luker’s cases. Some of these women’s testimonies seem to argue *against* the claim that subjects cannot critically examine their courses of action or develop preferences or desires which counter oppressive social norms once they have internalised the latter. Let me provide a brief example to justify this claim:

One of the interviewed subjects, for instance, acknowledges that “because of the way [she] was brought up”, visiting a doctor to be prescribed some form of contraception made her, at the time, feel uncomfortable. This woman reports that she did not want the doctor to have access to her personal information and medical records (which, presumably, would have made her “status” as a sexually active agent “publicly” available for that doctor and other health professionals).<sup>371</sup> However, looking back at her past decision, she deems her reason “silly”.<sup>372</sup> It seems then that this woman *endorsed* the values of her social milieu (those she had acquired as a result of her upbringing) and lived according to those values, but was later on able to assess them and even to (at least partially) reject those values.

It is worth noting that there is evidence of the fact that a case like the above is not an isolated one in Luker’s study. Indeed, Luker reports that her interviewees showed a tendency

---

<sup>371</sup> Luker, *Taking Chances*, 51.

<sup>372</sup> *Ibid.*

to retrospectively call into question the “reasonableness” of their past decisions.<sup>373</sup> Furthermore, while it is true that Luker claims that some women “felt trapped” by “social definitions around contraception”, it is also true that Luker frequently reports that her interviewees experienced some level of “conflict” when trying to navigate social imperatives and take into account their desires or preferences.<sup>374</sup> The latter description does not conclusively argue in favour of the internalisation of social norms in the sense defined above, which presupposed an impossibility to both see these norms as anything but “true” and to form desires or preferences which could diverge from the social norm (which seems to be necessary to experience some sort of “conflict”).

It seems then that the processes described above do not easily support the claim that agents who have “internalised” certain oppressive values are then *necessarily* incompetent in strong normative terms (as defined by Stoljar). Notably, it seems too quick to conclude that agents who are negatively affected by oppressive norms are so *irredeemably* – i.e. in a way that could neither be somewhat “reverted” by critical reflection nor at least become noticeable to agents themselves through motivations divergent from the oppressive social norms in question.

Admittedly, it could still be possible to defend the thesis that Luker’s agents lack *strong* normative competence even in light of the considerations above. Indeed, Stoljar could argue that one could call into question or revise a decision that was taken in light of oppressive norms without accessing “non-oppressive” or “true” social norms. For example, someone could question a sexist norm in light of another sexist norm. Imagine, for example, that one of Luker’s agents questions her past contraceptive risk-taking not because she is now able to perceive that the social norms that connect motherhood to self-worth are pernicious, but because she now realises that pregnancy was not a very effective way of gaining her partner’s

---

<sup>373</sup> Luker, “Contraceptive Risk Taking and Abortion,” 191; emphasis added.

<sup>374</sup> Luker, *Taking Chances*, 131.

recognition after all. In such a case, questioning the reasonableness of one's past decision would not amount to acquiring 'strong normative competence' in the sense defined above, because the agent could be questioning her decision *without* a grasp of the "right" norms. Or, in other words, agents could question a specific decision without challenging the underlying (oppressive) social norms that motivated that decision in the first place.

In reply, let me say that even if we decided against granting strong normative competence to agents who (like Luker's subjects) live under pernicious normative environments, it seems that we judge too quickly if we say that internalised oppression makes *full* and *permanent* endorsement of oppressive norms (and its derivative preferences) the *only* option available to oppressed subjects. My point is that the testimonies provided by Luker's agents and, indeed, the widely available personal stories involving resistance, self-questioning, conversions, and inner struggles with one's acquired values, seem to tell a different story.

In other words, it strikes me as an inaccurate description of gender oppression to say that the issue lies on women's radical incapacity to "perceive" norms as oppressive, harmful, distorting, and so on. To be sure, as I argued in Chapter 3, understanding and experiencing one's oppression might not be that straightforward and individual and collective work could be required in order to, e.g., take into account one's experiences of discomfort vis-à-vis one's oppression. However, the fact that noticing or conceptualising one's oppression is difficult is not equivalent to claiming that internalised oppression is *irreversible*.

Crucially, if the reader is still unconvinced, my (empirical) observations above can be backed up by turning to the theoretical resources presented in Chapter 2:

As Foucault noted, governing people implies a "versatile equilibrium [...] between techniques which assure coercion and processes through which the self is constructed or

modified by himself.”<sup>375</sup> According to Foucault, this involves, crucially, that subjection to power happens by way of both “external” and “internal” technologies. That is, a moral code (i.e. norms) is not internalized once and blindly or ideologically followed forever, it is insistently encouraged *externally* (e.g. by institutions, by the relationships we are in, by the “management” of our possibilities for action) *and internally* (e.g. by promoted forms of self-relations). Foucault indeed claims that “we have to get rid of the more or less Freudian schema”, i.e. “the schema of interiorization of the law by the self” on the grounds that this schema presents us with an oversimplified view of what it is to be governed in light of social guidelines of conduct.<sup>376</sup>

Reframing the discussion as suggested above has consequence for our debate: First, since external encouragement plays a role in the reinforcement of moral codes or social injunctions, then it might be possible to have more or less critical distance vis-à-vis the same social injunctions as our relationships, spheres of actions, and social circles vary. Second, since government by others also relies a great deal on repeated actions which are performed by governed agents themselves, there are a number of actions within the subjects’ reach which could potentially limit government by others. Let me briefly provide two key examples of the kinds of actions which are carried out (and *need* to be carried out) by governed agents themselves: (i) interpretation of the norms, and (ii) ethical self-constitution.

Regarding (i), I want to recall that, according to Foucault, subjects are not merely *made* subjects by norms, they also make themselves subjects and, in order to do the latter, they need to use, interpret, recreate, and navigate social norms in ways that are not fully-determined by these norms. For Foucault, individuals’ moral behaviours can coincide with what is prescribed by a ‘moral code’ but this coincidence is in no way a given.<sup>377</sup>

---

<sup>375</sup> Foucault, *About the beginning of the hermeneutics of the self*, 26.

<sup>376</sup> Ibid.

<sup>377</sup> Foucault, *The Use of Pleasure*, 25. Indeed, “morality” is for Foucault an ambiguous term. It refers both to the ‘moral codes’ defined in Chapter 2 and to morally relevant behaviour which can be confronted with the rules

Crucially, even if we were to focus on the “psychic” effects of social norms (as Stoljar does and as some psychoanalytic readings of Foucault do) the fact remains that, as Butler puts it, “power is not mechanically reproduced”.<sup>378</sup> Butler claims:

The psychic operation of the norm offers a more insidious route for regulatory power than explicit coercion, one whose success allows its tacit operation within the social. And yet, being psychic, the norm does not merely reinstate social power, it becomes formative and *vulnerable* in highly specific ways.<sup>379</sup>

In a word, the effects of oppression on the psychology of the oppressed, in a Foucauldian view, should not be thought as a one-off “inoculation” of fixed, inalterable, and irreversible oppressive normative prescriptions. Even if we saw norms as constituting subjects and, therefore, as “becoming a part of subjects”, it is not necessary to see (subjective) norms as “entities” which remain unaffected by the “use” that subjects make of them.

Regarding (ii), namely how government requires agents to perform actions on themselves to constitute themselves as subjects of a specific moral code, the Foucauldian distinction between morality and ethics introduced in Chapter 2 is pertinent here. Stoljar rejects oppressive social imperatives that could be thought as part of (what Foucault calls) the ‘moral code’.<sup>380</sup> Stoljar and Foucault could agree on the fact that agents cannot *choose* or easily change the moral codes available in a given context (indeed, Foucault sees moral codes as being much more stable than ethical self-relations).<sup>381</sup> Moreover, Stoljar and Foucault

---

and values determined by moral codes. Foucault disambiguates the term “morality” by referring to the first sense as “moral code” and to the second as the “morality of behaviours” [*moralité des comportements*].

<sup>378</sup> Butler, *The Psychic Life of Power*, 21.

<sup>379</sup> Ibid; emphasis added. See also Guillaume le Blanc, “Becoming a Subject in Relation to Norms,” in *Foucault and the Making of Subjects*. Le Blanc argues, in a way similar to Butler, that “there is a subjective life of norms” which is not entirely determined by a moral code. (Ibid., 134)

<sup>380</sup> Foucault, “On the Genealogy of Ethics,” 263.

<sup>381</sup> For example, Foucault claims that sexual prescriptions and prohibitions “did not change a great deal” if we compare the Greco-Roman moral codes, moral codes in early Christianity and, even, in “our societies”. Foucault identifies three main prohibitions: against excessive sexual activity, against extra-marital activity, and against homosexual activity. (Ibid., 265-6)

could agree that self-relations will be affected and directed by the prescriptions and prohibitions of a specific society.

However, Foucault would not infer from the above that individuals are irredeemably determined to be in a certain way. Moral codes do prescribe ethical relations, namely the kinds of relationships you ought to have to yourself when you act according to a specific moral code.<sup>382</sup> However, as I explained in Chapter 2, ethical relations are free and within our reach – i.e. even within a specific moral code it is possible to relate to ourselves *differently*.<sup>383</sup>

Moreover, Foucault's point is also that different self-relations can be achieved through the same 'techniques of the self' provided by the same code one may be trying to challenge. As it is well-known, Foucault turned to ethics in Ancient Greece late in his career to illustrate that other forms of ethical relations are possible. Crucially (to put it roughly) Foucault shows that ethics could be a matter of self-creation instead of being aimed at self-discovery, self-knowledge or self-renunciation.<sup>384</sup>

The significance of this example, as Allen notes, lies in showing that there is nothing inherently good or wrong about the specific 'techniques of the self' involved in ethical relations, "what matters is *how* they are used, to what ends and in what sorts of circumstances".<sup>385</sup> Take the technique of self-examination, for example, it is possible to self-examine in a way that does not contribute to one's subjection.<sup>386</sup> Indeed, the form of critical self-assessment that I have outlined throughout this thesis *is* a form of self-examination and self-assessment. Still, as I have argued, there is an important difference between carrying out self-examination with an understanding of how sociohistorical contingent factors

---

<sup>382</sup> Ibid., 263.

<sup>383</sup> Foucault, "The Ethics of the Concern of the Self as a Practice of Freedom," 284.

<sup>384</sup> Foucault, "On the Genealogy of Ethics," 265-6.

<sup>385</sup> Allen, *The Politics of Our Selves*, 63; emphasis added.

<sup>386</sup> Ibid.



conditioned the way one is *and* self-examining with an eye to making oneself conform to unquestioned social “truths” about oneself.<sup>387</sup>

I believe that the abovementioned points can even coexist with the worry that some subjects live in *pernicious* normative environments. That is: even if we were to accept the claim that some norms or stereotypes are objectively damaging, this does not force us to automatically attribute heteronomy to agents living in oppressive normative environments insofar as, some subjective strategies, may *limit* their oppression.

In the next section, I consider the challenges to procedural accounts of autonomy that may stand in spite of rejecting the ‘full-internalisation’ view.

## II. Beyond the “Full-Internalisation” Model:

In light of the discussion above, the following substantive objection could still be raised: even if internalisation does not cloud subjects’ reflective and normative capacities in a total and irreversible way, oppressed agents still live in an (objectively) harmful normative world. That is, regardless of the fact that the effects of oppression on the psychologies of those oppressed could be limited, reverted, or circumvented by those oppressed, the fact remains that these subjective strategies do not change the essential: oppressive norms will still be “out there”.

One strategy out of this impasse could be to claim that, the subjective strategies which make agents *less* subjected to particular oppressive moral codes, are effective enough to be considered autonomy-enhancing practices, even if they are not sufficiently “world changing”. After all, as I have already argued, I am interested in a model of autonomy that does not presuppose the (unrealistic) possibility of creating one’s own norms but rather that explores

---

<sup>387</sup> See also Beatrice Han-Pile, “Foucault, normativity and critique as a practice of the self,” *Continental Philosophy Review* 49, no. 1 (2016): 99.

the different ways in which agents can relate to already existing social norms. The “transcendence” of social norms (i.e. the fact that they exist beyond agents’ wills or their strategies to relate to or cope with those norms) is always the reality autonomy theorists need to work with. The position that these norms could be “objectively” good or bad does not significantly alter, I believe, the fact that agents will always have limited control vis-à-vis those norms.

I start this section by considering an example which illustrates what I have in mind when I claim that it might be possible to relate to the same oppressive norms in different ways and even to “play” with the norms available in an oppressive context. In light of this example, I consider: (i) the advantages of holding on to a procedural theory of autonomy to deal with cases like this; and (ii) whether there is still a point to be made against proceduralism even conceding that oppressive norms enable some “room for play”. Analysing this second point requires us to clarify what exactly is gained in oppressive contexts (e.g. what can/cannot be changed or reversed) through subjective strategies like the ones I describe through my example.

Consider the case of the Brazilian Black trans-gender artist Rosa Luz. First, it is worth noting that Luz’ story is in itself a case against the thesis which I problematized in the previous section (i.e. on the impossibility to critically assess norms after they have been internalized): Luz was educated in a conservative religious environment and manifests that she felt comfortable in this environment (e.g. she was one of the best students in her Catholic school and even jokes that, had she been a cis heterosexual man, she would have turned out to be an ultra-conservative).<sup>388</sup> According to Luz, when she fully acknowledged her trans identity, she was presented with the dilemma that her identity was incompatible with her

---

<sup>388</sup> See Gabriel Galli, “Em país hostil a travestis, brasileira faz arte a partir de sua identidade e denuncia transfobia,” *Global Voices*, July 6, 2020, <https://pt.globalvoices.org/2020/07/06/em-pais-hostil-a-travestis-brasileira-faz-arte-a-partir-de-sua-identidade-e-denuncia-transfobia/> (last accessed September 2020)

religious convictions and eventually dropped the latter. But Luz' personal development is not my focus here.

In the documentary *Enough with Catcalling*, Luz is shown during a performance in Brasilia: she is dressed in “feminine” clothes (she is wearing a skirt, makeup, high-heels) and takes her shirt off in the stairs of a station.<sup>389</sup> Leaving aside the many aims achieved by her performance (e.g. giving visibility to trans bodies, which is a key theme in her work), I want to reflect on the clever use that Luz made of available gender norms, of available stereotypes and, even, of Brazilian law. As it happens in many other places, appearing topless in public spaces is typically deemed (socially and legally) acceptable for ‘men’ but not for ‘women’.<sup>390</sup> By standing there, Luz presented passers-by and the police with the following dilemma: forcing her to leave the station (e.g. by showing hostility to her or by straightforwardly arresting her) would have meant acknowledging that she *is* a woman; on the other hand, allowing her to stand there could be seen as a disregard of her self-identification as a woman but gave her, nonetheless, a “victory” of a different kind – she could protest and make her cause visible. Luz acknowledges that her body “falls through the cracks” of Brazilian gender norms and laws and reports that she is then determined to make her body visible “through those cracks”.<sup>391</sup>

Admittedly, acts of resistance like this do not (immediately) change the normative environment itself (i.e. oppressive gender norms or patriarchal and transphobic laws will still be there at the end of the day), but I believe that it is not hard to see how Luz' creative use of the available norms and laws contributes to ends that are *not* merely those of *reinforcing* her oppression.

---

<sup>389</sup> *Chega de Fiu Fiu (Enough with Catcalling)*, directed by Amanda Kamanchek and Fernanda Frazão (2018; Brazil: Brodagem filmes). The performance can also be watched at <https://pt.globalvoices.org/2020/07/06/em-pais-hostil-a-travestis-brasileira-faz-arte-a-partir-de-sua-identidade-e-denuncia-transfobia/>

<sup>390</sup> I say ‘typically’ because an obvious exception would be the hyper-sexualised representation and appearance of women in public spaces that is culturally allowed and encouraged (e.g. during carnival).

<sup>391</sup> *Ibid.*, min 10:47; my translation.

To be clear, I am not suggesting that resistance via a creative or non-orthodox use of the available norms should necessarily involve as elaborate or thought-through actions as the example above suggests. To the contrary, as I argue in Chapter 5, I believe that it is even possible to engage in rule and role “bending” by accident, at least in a first instance.<sup>392</sup> The point that I want to convey for now is merely that agents living in normatively pernicious environments could become *less* heteronomous by working with the “corrupted” materials available to them. Moreover, I want to highlight that this is the kind of subjective strategy that could be overshadowed when we rule out autonomy by making strict reference to the “quality” of socially available norms or stereotypes.

The question that should be asked now, however, is whether the possibility of agents adopting strategies like Rosa Luz’ is enough to disregard substantive theorists’ worries like the ones discussed above. I consider below whether there is still a point to be made *against* proceduralism *even* if we concede that available oppressive norms could leave room for play to *limit* one’s oppression.

To begin with, Stoljar could argue that even granting that the possibility of relating more critically to pernicious social norms exists, her key substantive point still holds because we could not distinguish between emancipatory and oppression-reinforcing ways of executing the same *formal* agentic skills without attributing a *value* to certain social norms and conditions. Stoljar can still argue that we need to take a stance on the value of the norms and social conditions *themselves* before assessing how autonomy-undermining or autonomy-enhancing individuals’ strategies to live or cope with those norms are. Consequently, some attitudes could not be deemed “autonomy-undermining” unless autonomy theorists

---

<sup>392</sup> As I argue in Chapter 5, accidental events, e.g. the accidental enactment of different social roles, could indeed help *see* available norms otherwise or to experience discomfort or alienation in a way that was not possible before. I do not suggest that no reflection or elaboration is necessary, but rather that the materials for reflection and critical self-assessment do not necessarily become available in a reflective or intellectual manner.

renounce their commitment to providing a “morally neutral” test of autonomy.<sup>393</sup> The latter amounts to renouncing a proceduralist project and endorsing a substantive one instead.

Concerning this point, let me say that I agree with the fact that *some* substantive normative commitments might have to be adopted (I say more about this towards the end of this chapter), but I wish to give much more credit than Stoljar does to the potential that procedural strategies may have in some oppressive normative environments. Even if we had an “objective” account of what counts as a pernicious or oppressive social norm, a radical substantive solution like the one proposed above would obscure the ways in which different forms of self-government might limit the effects of the *same* pernicious norms. There is the danger that on such a radical substantive solution we would basically have to condemn most past and existing contexts as incompatible with autonomy in a blanket way, not differentiating between them or between how different agents navigate them.

Moreover, as I said in previous chapters, the move of *automatically* considering agents heteronomous when they are subjected to “oppressive” norms and stereotypes seems to be more in line with an ideal of ‘orthonomy’ (i.e. of government by right norms) than with one of *self-government*. Crucially, it is worth noting that an ideal like ‘orthonomy’ does not necessarily succeed in *limiting heteronomy*, insofar as to do so we would need to address the epistemic question (and the related ethical and political questions) of who could legitimately claim to know what the objective goods are, how we would know who these would be, and what authority they should be given. Depending on one’s views on the possibility of success of the project of establishing what is objectively good, one might think that making accounts of autonomy dependent on this success, is unreasonably making autonomy models hostages to fortune in meta-ethics and metaphysics.

---

<sup>393</sup> Stoljar, “Autonomy and adaptive preference formation,” 232.

Nonetheless, as I anticipated in the beginning of this chapter, there are at least two strategies to make a substantive critique *without* the need to neither committing to the model of full-internalisation that I have problematized above nor overlooking the room for play that oppressed agents may have in oppressive normative environments:

First, we could look into the psychological effects of oppression *irrespective* of internalisation – e.g. by claiming that the harm done to the psychologies of the oppressed is not tied to the damage done to their ‘normative competence’ (i.e. to their capacity to perceive certain norms as oppressive, damaging or false) but to the mere *fact* that agents affected by oppressive norms or stereotypes do not have the choice to ignore these norms. So, for example, one could say that even if I am aware of the falseness of some of the social stereotypes that apply to me, or critical enough to “play” with them, the fact that I am conditioned and constrained by these norms and stereotypes remains.

To put it plainly, the fact for Stoljar seems to be that I do *not* get to *choose* whether to direct my attention, my capacities, and energies to deal with oppressive stereotypes. Rather, my participation in society will *require* me to do all the latter. Moreover, Stoljar’s point is that the psychological life of the oppressed is not as free as that of those who are not oppressed by the same norms and stereotypes and, therefore, do not have to respond to them. I develop this position in section *a*, by taking a look at Stoljar’s more recent work.

Second, we could look at the damage done by oppression more broadly, i.e. beyond its psychological harms, and focus on the way in which it affects individuals’ external material conditions, e.g. the social relations in which they live. As I mentioned earlier in this chapter, this is Oshana’s strategy. I consider this strategy in section *b*.

Both Stoljar and Oshana share the conviction that *even* those agents who attempt to *resist* their oppression (like Rosa Luz) are *heteronomous* by virtue of the external conditions (be them normative or relational) that oppress them.

a) Psychological oppression beyond the “full-internalisation” model:

In a more recent article, *Living Constantly at Tiptoe Stance*, Stoljar argues that the interrelation between an oppressive context and the psychologies of those oppressed is a very complex one and should not be reduced to the issue of whether the oppressed actually endorse or accept oppressive norms.<sup>394</sup>

Stoljar analyses cases in which an oppressive normative context (e.g. one marked by racist stereotypes) affects the lives and psychologies of those oppressed *even* when they *resist* or *explicitly reject* the negative stereotypes of their contexts. Stoljar claims that, living in a normatively pernicious environment, means having one’s self-conception and the available spaces of possibility for one’s life and actions limited by negative “social scripts”.<sup>395</sup> Stoljar provides several examples of the above but, for the sake of brevity and clarity, I turn to the one that does not require an extensive introduction:

Stoljar reflects on the famous case of Martin Luther King Jr. His incontestable role as a political activist and as an advocate of Black people’s rights can be taken as conclusive evidence of the fact that he both *rejected* and actively *resisted* the oppressive social meanings in light of which Black people in the USA (and, certainly, in many other places too) had to live their lives.

Still, Stoljar argues that King’s lack of endorsement of the oppressive normative context in which he lived does not amount to ‘psychological freedom’ because, as King himself wrote, he (and others like him) were “living constantly at tiptoe stance, never quite knowing what to expect next, (...) plagued with inner fears and outer resentments” and “forever fighting a degenerating sense of ‘nobodiness’”.<sup>396</sup> Then, *regardless* of the fact that

---

<sup>394</sup> Natalie Stoljar, “Living Constantly at Tiptoe Stance: Social Scripts, Psychological Freedom, and Autonomy,” in *Personal Autonomy and Social Oppression*, ed. Marina Oshana (New York: Routledge, 2014).

<sup>395</sup> *Ibid.*, 106.

<sup>396</sup> Martin Luther King in *Ibid.*, 113

King rejected the pernicious social scripts that were thrust upon him as a Black person, his psychological life was *in fact* scripted by these stereotypes to a significant extent.

An important thing to notice is that Stoljar's point does not just apply to cases of oppression that are sustained because of (explicit) severe legal constraints. Some of her other examples show, for instance, that the situation of Black people in the United States is *still* negatively scripted even after the civil-rights movement. Stoljar mentions, for example, how black men need to explicitly disavow the racist stereotype of black men frequently being criminals if they, say, approach someone on a street at night.<sup>397</sup>

Stoljar argues that social scripts interfere with psychological freedom, even without endorsement of the social scripts by the agent, because people affected and oppressed by negative scripts are forced to take this scripting into account in several ways: they need to “*adapt* their decision making and behavior to it”; they are still “called on to *take an evaluative stance* to the norms embedded in the script”; and they “must *respond to and disavow* the script to see themselves and be seen as equal participants in their daily transactions with others”.<sup>398</sup>

Gender oppression can also be thought as a form of “oppressive scripting” and, therefore, Stoljar's points are also useful to analyse gender oppression. Indeed, other authors make an analysis similar to Stoljar's when they reflect on the effects of the sexual objectification of women in contemporary societies. For example, Timo Jütten claims that sexual objectification is harmful to women's autonomy both because (i) a public status of ‘sexual object’ is *imposed* on women (and, thus, women are denied status as *self*-presenters, i.e. as individuals who have power over the constitution of their public self-image), and (ii) because being sexually objectified means having the socially available meanings for public self-presentation and participation in social interaction severely limited.<sup>399</sup> So, like Stoljar's,

---

<sup>397</sup> Elizabeth Anderson in *Ibid.*, 119.

<sup>398</sup> *Ibid.*, 107.

<sup>399</sup> Timo Jütten, “Sexual Objectification,” *Ethics*, 127 (October 2016): 34-5. Jütten adopts the notion of “self-presentation” from Thomas Nagel's and David Velleman's works. “Self-presentation” refers to the



Jütten's position highlights that social stereotyping has effects on one's self-definition and self-presentation regardless of the affected agents' stances on these stereotypes.

Stoljar takes the "involuntary" manifestation of the social script in one's self-conception as evidence of the "internalization" of the negative social script and, consequently, as evidence of oppressed agents' *undermined* psychological freedom and autonomy.<sup>400</sup> The notion of 'internalisation' at work here is, however, different from the one that was presented in the previous section. Stoljar claims:

I suggest, [...], that there is a broader notion of internalization that does not require endorsement. Oppressive social scripts are embedded in agents' psychologies in ways that call for multiple psychological responses—for instance, anticipation, adjustment, accommodation, and evaluation—that are not typically required from agents to whom neutral or "dominant" scripts apply. Agents subject to oppressive scripts are "constantly at tiptoe stance."<sup>401</sup>

Therefore, whether an agent has internalised an oppressive social environment is now decided in terms of whether that social environment *manifests* in her psychology in the ways listed in the paragraph just quoted. That is, internalisation can be "partial" and is not incompatible with *rejecting* oppressive norms or stereotypes.<sup>402</sup>

To sum up, Stoljar implies that agents who have internalised their oppression (in the new sense defined just above) lack 'psychological freedom' because their psychologies are shaped by their oppression in a way that is outside their control. Moreover, this lack of control over one's psychological life is not one that could be gained (or regained) by critically

---

phenomenon by which "all of us routinely create public images of ourselves by selecting which aspects of ourselves we expose to others and which we conceal from them." (Ibid., 33) Social interaction requires that "others can recognize our public image as an intentional presentation of ourselves" (Ibid.) and, therefore, "not being recognized as a self-presenter is an existential threat to our agency" (Ibid., 34).

<sup>400</sup> Stoljar, "Living Constantly at Tiptoe Stance," 106.

<sup>401</sup> Ibid., 118.

<sup>402</sup> Ibid.

assessing one's context or by intellectually rejecting it, regardless of whether these critical/reflective tasks are possible. On these grounds, Stoljar concludes that procedural accounts work with an insufficient account of 'psychological freedom' (i.e. as competence, and absence of pathology and coercion) and she argues that an agent's psychological freedom depends on "social features of the world".<sup>403</sup>

From the above, the (weaker) substantive thesis  $S_2$  follows. (To recall: A substantive account of autonomy is the only viable approach under conditions of 'subjection' because oppressive norms or stereotypes will limit one's psychological freedom, and therefore, one's autonomy, *regardless* of one's stance on them.)

Are the concerns vis-à-vis substantive accounts which I raised in the previous section still relevant given this revised account of internalisation? I argue that they are and that, indeed, some of my worries could be reframed more clearly in light of Stoljar's revised position:

First, I reiterate the worry that Stoljar's account does not allow us to distinguish (for the purposes of autonomy) between agents who manage to significantly limit the effects of oppressive norms on themselves and those who do not at all. Nonetheless, the signs of "partial internalisation" that Stoljar rightly identifies (e.g. one's ability to reject oppressive norms while still being affected or conditioned by them) could be seen as steps towards being *less* oppressed. If, to put it in the Foucauldian terms introduced earlier, governing people relies on external *and* internal strategies to direct one's conduct, then being able to see social injunctions as false, contingent or pernicious (and not as ultimate "truths" about oneself) means at least that one of the dimensions of government (i.e. the internal one) has been made less effective.

---

<sup>403</sup> Ibid., 106.

As Stoljar notes, her account is indeed well-suited to deal with cases of agents who “have practical control over their daily routines” but are, nonetheless, subjected to pernicious social injunctions on who they are or how to be.<sup>404</sup> However, Stoljar’s model is missing resources to account for the different forms in which agents *with* practical control over their daily lives could also affect the *degree* to which social injunctions constrain their lives and self-definitions.

Acknowledging that agents’ autonomy may come in degrees instead of being an all or nothing matter may have an additional advantage in this case, namely that of providing a more realistic account of ‘psychological freedom’ than Stoljar does here. What I mean is that, even conceding that agents affected by oppressive social scripts cannot choose to ignore these scripts (as Stoljar notes, they will have to anticipate and adjust their behaviours in light of these scripts), I believe that the same applies, in many ways, to most agents and perhaps to all. Any process of socialisation implies working with social scripts that one cannot fully control and that one cannot just ignore.

For example, gender norms in patriarchal contexts also limit spaces of possibility for ‘men’. As Meyers notes, “masculine socialisation is powerfully directive” too – e.g. “little boys are taunted for “sissyish” behaviour” and guided into certain social roles through the male models with whom they identify.<sup>405</sup> Moreover, as I have argued, one consequence of challenging the picture of oppression presupposed by the models of manipulation and deception is that it allows us to question the idea that oppression is due to the intentional actions of specific individuals – this is not how *structural* oppression works. Consequently, we should avoid seeing social power relations as relations involving, on the side of the “oppressors”, transparent intentions and absolute freedom from social injunctions and, on the side of the “oppressed”, self-deception and scripted lives. Both parties play social roles

---

<sup>404</sup> Ibid., 112.

<sup>405</sup> Meyers, *Being Yourself*, 7.

that transcend their *subjective* intentions. As a result, socialisation *always* necessarily entails (i.e. even if one's life is not scripted by *oppressive* stereotypes) having to navigate *social* scripts.

Admittedly, Jütten's point (sketched above) could be used to signal an important objection to my reply. We could concede that all agents are exposed to (and conditioned by) social norms and stereotypes in a way that is not entirely within their control and yet argue that some agents will be *especially* damaged by the norms and stereotypes that affect them. Namely, because of some *oppressive* social meanings and stereotypes, some agents will be damaged in their capacity as self-presenting agents, which includes (broadly speaking) one's ability to "influence how others see us".<sup>406</sup> A diminished capacity as a self-presenter harms agency in a fundamental way, insofar as (according to Jütten) this harm threatens "one's ability to make claims on one's own behalf".<sup>407</sup>

Might the above be a way in which oppressive social stereotypes and norms affect those oppressed by those norms more deeply and seriously than the rest? Certainly, and this is, I insist, why distinguishing degrees of autonomy makes sense. Importantly, if the worry is that some oppressive stereotypes might be so pervasive that they might radically restrict what oppressed agents can *do* with their claims and actions in their social world, then I would reply that cases like these could be captured by my 'non-domination' condition (presented in the final section of this Chapter and developed in Chapter 5).<sup>408</sup> As I anticipated, I give a separate treatment to those oppressive cases in which agents are "stuck" in a position which is non-

---

<sup>406</sup> Jütten, "Sexual Objectification," 34.

<sup>407</sup> Ibid. This point can be related with my earlier point about 'strong normative competency' in Benson. What is crucially damaging for the agencies of those oppressed by social norms and stereotypes is that they might lose control over the "self-disclosing" dimensions of their actions. That is: personal or moral features might be attributed to them independently of their actions.

<sup>408</sup> One could think of (some) feminists' arguments against pornography as building up on this kind of worry. For example, one could argue that the way in which heterosexual intercourse is typically represented in pornography results in a sexist stereotype which damages women's agency in a fundamental way. Namely, pornography perpetuates the idea that "all women, at all times, whether they admit it or not, 'really' want sex, and that their refusals of sex are not sincere or authentic; women say 'no' but mean 'yes'." (Lorna Finlayson, "How to Screw Things with Words," *Hypatia* 29, no. 4 (Fall 2014): 777) Crucially, the "silencing effect" caused by pornography would affect what women can *do* with their words. As Finlayson explains: "The silencing performed by pornography is "illocutionary" in a double sense: it is an illocutionary act of silencing; and it silences women at the illocutionary level, by preventing them from doing things with their words." (Ibid.)

reversible through the margin of freedom available to them until external change happens. In the extreme hypothetical case imagined here, symbolic cultural change would be a precondition of agents having a meaningful margin of freedom available to limit their heteronomy.

My point for now is, nonetheless, that in cases where a significant room for play *is* available, we are not necessarily giving the oppressed more of a voice by downplaying their efforts to make themselves less heteronomous within their oppressive environments. Just the opposite, one way to acknowledge the deeper and more serious nature of the effects of oppressive norms on them, is to differentiate the levels of heteronomy that exist among them and to recognise, indeed celebrate, the reduction of heteronomy some of them are still managing to achieve. What is more, insofar as these efforts may frequently involve acts of self-definition or self-identification (e.g. as “fighters”, ‘Civil Rights Activist’) we should even consider if, by overlooking oppressed agents’ efforts to make themselves less heteronomous, we might not even be damaging their status as self-presenters further.

To summarize, if we consider Rosa Luz’ case again, it is not impossible to claim that, even though her self-presentation is not entirely up to her, and even if her appearance in a public space is conditioned by the stereotypes on women and on transgender people that the spectators of Luz’ performance hold, she nonetheless managed to use the stereotypes of her spectators in her favour. In a way, *who* others see is not entirely within Luz control but, by playing with available norms and stereotypes, Luz unveils that her self-presentation is not entirely within the control of those in a dominant position either.

Before concluding, we need to consider an additional way in which the psychology of the oppressed might be harmed or burdened by oppression. As I briefly explained in the introduction to this chapter, Benson analyses cases in which an oppressive socialisation (and the oppressive stereotypes and norms in that oppressive environment) affects how

individuals see themselves and relate to themselves. Locating Benson's strategy vis-à-vis the other strategies analysed in this chapter is not easy:

On the one hand (as my discussion in part I acknowledged), it is implied by some of Benson's remarks that those who have internalised problematic attitudes towards themselves as a result of an oppressive socialisation are in a situation which appears irreversible (or extremely difficult to reverse) through subjective means. For example, an oppressive socialisation, Benson claims, "systematically prevents many women from recognizing more adequate views of their real strength and value".<sup>409</sup> Indeed, as I mentioned above, Benson is sceptical about the utility of procedural historical tests in cases where agents have internalised negative self-images insofar as an internalised negative self-image might increase one's tolerance to illegitimate past interferences.<sup>410</sup>

On the other hand, Benson resists accounts of female oppression which, like Stoljar's, conflate self-government with orthonomy. A model of autonomy, Benson argues, should accommodate "arenas for some autonomy within the evaluative, psychological, and political frameworks that undergird relations of domination and subordination."<sup>411</sup> For example, Benson is willing to concede some degree of autonomy to some of Luker's agents, which suggests that, in his view, there *are* ways in which oppressed agents can make themselves differently (and, notably, *less*) affected by the *same* oppressive stereotypes.<sup>412</sup>

I argue that the abovementioned ambivalence is reflected in the kinds of examples that Benson provides. Some of Benson's examples could be seen as cases of 'domination' and, therefore, as not meaningfully reversible through subjective strategies alone. Other cases offered by Benson, however, could qualify as forms of 'subjection' and, as such, could be

---

<sup>409</sup> Benson, "Autonomy and Oppressive Socialisation," *Social Theory and Practice* 17, no. 3 (Fall 1991): 396.

<sup>410</sup> *Ibid.*, 394.

<sup>411</sup> Benson, "Feminist Intuitions and the Normative Substance of Autonomy," 130.

<sup>412</sup> *Ibid.*, 128. See footnote 362

significantly limited through a (revised) procedural account like the one I propose in Chapter 5. I justify these claims below:

Benson developed his critique of content-neutral procedural accounts by elaborating on the plot from the movie (and play) *Gaslight*. Roughly, this movie portrays a man deliberately trying to convince his wife that she is mentally unstable: he confuses and disorients her, manages to isolate her from other people, makes her doubt her memory and her reasoning capacities, among other things.<sup>413</sup> As a result of the husband's actions, the woman's sense of her own self-worth is completely eroded.

Admittedly, a case like the one described above could hardly be thought compatible with procedural independence insofar as it could be deemed a case of straightforward manipulation. However, Benson's point is to show that the 'gaslighting' model can do without the dyadic and subjective-intentional elements which make it a case of manipulation: Benson asks us to consider cases where a person develops a sense of "incompetence and estrangement from her conduct on the basis of reasons that are accepted by a scientific establishment which is socially validated and which she trusts".<sup>414</sup> Some forms of feminine socialisation would be, according to Benson, compatible with this description.

"Social gaslighting", as described above, would qualify as 'subjection'. Indeed, social gaslighting appears to share many features with the forms of non-disruptive oppression which I have analysed in previous chapters: social gaslighting involves oppression through socially-accepted "truths" which classify, define, and shape agents throughout their "normal" developments; these socially-accepted "truths" are typically accepted by agents themselves and taken into account in self-relations; and the whole process cannot be explained exclusively by the actions of identifiable deceivers or manipulators. Is a procedural strategy completely futile in cases like this?

---

<sup>413</sup> Paul Benson, "Free Agency and Self-Worth." *The Journal of Philosophy* 91, no. 12 (December 1994): 655.

<sup>414</sup> *Ibid.*, 657

I argue it is not. Indeed, as I proposed in Chapter 3, one of the advantages of the revised procedural model that I propose is that it might help agents relate differently to themselves and, crucially, to the negative feelings they might have towards themselves as a result of their oppression. To recall, I have argued that one of the advantages of carrying out self-assessment while informed about the perspectives of others who occupy social positions similar to mine (e.g. who are affected by the same social stereotypes) is that this might help agents push back on feelings of shame, worthlessness or failure. For example, I argued that sharing experiences of discomfort with others (e.g. with those also struggling to meet the demands of a social role) might help agents see that their suffering is not connected to a lack of self-worth but, say, to the unreasonable or pernicious expectations of a social role.

However, it is necessary to note that some of the cases that Benson refers to involve agents who are “crushed” by their oppression. These latter cases, I argue, should be addressed differently. For example, Benson mentions that (social) shame can lead to “a collapse of the person’s sense of worthiness to act”.<sup>415</sup> According to Benson’s interpretation, this can be one of the effects of slavery, Benson claims: “[t]he social mechanisms that sustained slavery did so in part by working to destroy slaves’ sense of their competence to make their own decisions and manage their own lives”.<sup>416</sup>

In these latter cases, I argue, there is something different from ‘subjection’ going on. The cases of the slave and of the woman gaslighted by an abusive husband, involve radical asymmetries in power relations, crushed senses of agency, and intentional acts aimed at annihilating (rather than “conducting”) the agency of the oppressed. Indeed, Benson emphasizes that agents in these more radical cases of gaslighting are particularly ill-suited to participate in relationships and interactions with others –<sup>417</sup> i.e. what is particularly bad in

---

<sup>415</sup> Ibid., 658.

<sup>416</sup> Ibid., 659.

<sup>417</sup> Ibid., 660.



these cases is not exclusively the role that individuals play but the limited room for play that they have in social relations which are frozen and profoundly asymmetrical. In other words, it seems that these latter cases qualify as cases of ‘domination’ (in the broad sense defined in Chapter 2).

To conclude, I believe that the model I propose in Chapter 5 provides tools to deal with the different forms of ‘gaslighting’ proposed by Benson.

b) Non-psychological harms of oppression:

In this section, I consider Oshana’s view. As I anticipated in the introduction to this chapter, Oshana defends the idea that an agent’s psychology should *not* play a definitive role in determining her autonomy. Indeed, Oshana is critical of the “largely subjective” character of accounts of autonomy which decide an agent’s autonomy based on a combination of “the structural and developmental character of her judgments, preferences, and pro-attitudes” *and* the agents’ “attitude towards these”.<sup>418</sup>

But what is exactly under criticism when Oshana refers to the “largely *subjective*” focus of procedural accounts? First, Oshana notes that, for procedural theorists, the psychological state that defines autonomy (i.e. one’s psychological structure and one’s attitude vis-à-vis that structure) is defined and identified by the agent whose autonomy is the question. Therefore, in procedural accounts, “the adequacy of a person’s psychological state [...] is not a matter for others to decide” or, in other words, “only individuals can be the measure of their sense of autonomy or feeling of autonomy”.<sup>419</sup>

Oshana believes that the latter is problematic because this solution (wrongly) conflates the subjective *sensation* or feeling of being autonomous with the objective *fact* of being autonomous. Oshana claims: “sensation and feeling are qualia, and qualia alone do not decide

---

<sup>418</sup> Oshana, *Personal Autonomy in Society*, 50; emphasis added.

<sup>419</sup> Ibid.

the *fact* of autonomy any more than a feeling of oneself as nonautonomous decides against autonomy”.<sup>420</sup> So, in Oshana’s view, a “happy slave” is heteronomous by virtue of his social status as a slave, his subjective assessment of his own situation does not make him autonomous. Crucially, a slave would be heteronomous irrespective of the degree to which she may have come to internalise her oppression.

Another reason why, according to Oshana, the subjective focus of procedural accounts is inadequate relates to the *aims* that Oshana sets for her model. That is, in addition to deeming the subjective strategy of procedural accounts *inaccurate* (because it mistakes “feelings” of autonomy for actual autonomous states), Oshana deems subjective approaches *ineffective* in light of the aim of enhancing or empowering agency.

Oshana is interested in the possibilities and limits of “self-determination under conditions of oppression.”<sup>421</sup> Having this aim in mind, she deems subjective strategies ill-suited to ensure self-determination because, in her words:

... psychological autonomy of a sort favoured by content-neutral theorists is incapable of empowering a person in many areas of life that are crucial to successful agency where conditions of oppression are present or probable.<sup>422</sup>

Oshana’s critique of procedural accounts is thus not only raised on the grounds that procedural accounts are ill-suited to make an appropriate *diagnosis* of people’s heteronomy. Additionally, Oshana claims that procedural accounts are ill-equipped to *enhance* the autonomy of those who are oppressed.

---

<sup>420</sup> Ibid., 50-1; emphasis added.

<sup>421</sup> Marina Oshana, “Is Social-Relational Autonomy a Plausible Ideal?,” in *Personal Autonomy and Social Oppression*, 5.

<sup>422</sup> Ibid.

Crucially, Oshana suggests that it is not possible, through exclusively subjective means, to transform “a situation of constraint into one of independence”.<sup>423</sup> Focusing too much on psychological strategies to tackle oppression (e.g. on changing one’s self-reflective attitudes or on “filtering” the social values that one endorses), might end up downplaying the many other challenges that the oppressed face when trying to live their lives in a self-determining way. Alternatively, Oshana argues that “persons [...] fail to be autonomous because they lack characteristics that only a *social* theory of self-determination can supply”.<sup>424</sup>

In light of all the above, Oshana proposes to define autonomy by reference to agents’ *social* status rather than by their *mental* status. Oshana’s model includes some conditions that are traditionally associated with procedural models (i.e. epistemic competence, rationality, procedural independence) but also other conditions that look beyond agents’ psychologies. Namely, according to Oshana, an autonomous agent should be in a position where self-respect, control, access to a range of relevant options, and substantive independence are secured.<sup>425</sup>

I do not go back to the conditions which are typically shared with procedural accounts because they are defined in a way sufficiently similar to the procedural accounts that I have analysed in previous chapters. Likewise, self-respect and substantive independence can be understood in line with the presentation provided earlier in this chapter. Let me then unpack Oshana’s distinctive account of self-determination by paying special attention to the conditions of ‘control’ and ‘access to a range of relevant options’. In particular, I am interested in Oshana’s attempts to secure self-determination in the face of oppression or of the possibility thereof.

---

<sup>423</sup> Oshana, *Personal Autonomy in Society*, 89.

<sup>424</sup> *Ibid.*, 52; emphasis added.

<sup>425</sup> *Ibid.*, 76-86.

Oshana's 'control' condition is one of the ways in which self-determination is safeguarded. Oshana claims that an autonomous agent should have the power to "determine how she shall live"<sup>426</sup> and she unpacks this claim as follows:

... a person who is autonomous must enjoy a social status that safeguards her right to manage key aspects of her life against other persons or institutions that might attempt to wield coercive control over her. The autonomous person must be empowered to challenge others who might attempt to direct her against her wishes, or who might aim to dominate her, even if she never has cause to do this.<sup>427</sup>

Through this control condition, then, Oshana grants that agents are sufficiently safe from coercion, both actual and potential. I analyse this claim in steps:

First, how to understand coercion is not entirely clear from the abovementioned quote. Is 'coercion' just any *unwanted* intervention? This seems suggested by Oshana's claim that the autonomous person must have sufficient control over relevant aspects of her life so as to be able to limit any attempt to "direct her against her wishes".<sup>428</sup> Such an interpretation, however, would not be entirely compatible with Oshana's general strategy, insofar as resorting merely to the agent's value framework (e.g. to what *she* judges to be acceptable interventions), would bring us too close to the "overly subjective" account of autonomy that Oshana rejects.

Another way to read the passage above is possible: We could say that there *are* limits to what kinds of interventions are acceptable independently of what agents could consent – namely, no intervention should result in the agent's *de facto* heteronomy. This is not only compatible with Oshana's general framework but also seems to be suggested in the quoted

---

<sup>426</sup> Ibid., 83

<sup>427</sup> Ibid., 84

<sup>428</sup> Ibid.

passage: Oshana claims that an autonomous agent must be empowered to challenge (in addition to others' attempts to conduct her against her wishes) others' attempts to "dominate her".<sup>429</sup> Therefore, regardless of what an agent feels or wishes, interventions that result (or could result) in domination could emerge as a standard of "objectively" illegitimate interventions, in the sense that domination is a state in which agents *de facto* lose control over relevant aspects of their lives. So, autonomous agents must be both protected from actual domination and empowered to limit attempts of domination.

The claim that control requires a certain power over actual and likely interventions takes me to my next point. According to Oshana, the kind of control required for autonomy should not only be factual but also *counterfactual*. That is, an agent lacks the kind of control over her life needed for self-determination (and, therefore, for autonomy) not only when she is coerced, but also when coercion is *likely* or *possible*.<sup>430</sup> Oshana considers the following case:

We cannot claim that a person is self-governing if her efforts to determine how she will live would have been thwarted had she tried to act differently with respect to activities relevant to the direction of her life".<sup>431</sup>

To illustrate this point let me quickly go back to the case of Kaew, presented in Chapter 3. To recall, Kaew had been a victim of sex trafficking but ended up willingly running the "bar" in which she worked, even after her "debt" with her exploiters had been paid. For the sake of clarity, let's consider Kaew's situation only from the moment onwards when she has *accepted* her condition as a sex worker and decides to stay and work in the bar. In Oshana's model, Kaew would be heteronomous even after coming to terms with her life as a sex worker. The reason for the latter is that, had she not changed her mind or adapted her

---

<sup>429</sup> Ibid.

<sup>430</sup> Ibid., 83

<sup>431</sup> Ibid., 84

preferences (e.g. if she had tried to escape), she would have been most likely *forced* to stay by the people who put her in that situation in the first place.

Additionally, Oshana considers ‘access to a range of relevant options’ a further condition for self-determination. In line with the strategy described above, what counts as ‘relevant’ is not defined by the subject whose autonomy is the question. Oshana claims:

[d]eliberate acceptance of a particular state of affairs as one that offers an adequate range of choices does not prove autonomy any more than repudiation or disavowal of the same state of affairs establishes a lack of autonomy. [...] the fact that a person finds her social situation acceptable does not mean that hers is an acceptable situation.<sup>432</sup>

So how are ‘adequate’ options and ‘relevant’ aspects of a person’s life defined? Oshana claims that self-determination is possible when she retains control over domains which are “significant for the person’s agency”.<sup>433</sup> Example of these areas of significance where options must be available include “choice of partner, the decision to have children, employment, religious expression, and so on.”<sup>434</sup>

All in all, I want to retain two key elements from Oshana’s account:

First, it is necessary to deem cases in which there is no margin for any “meaningful” or “significant” freedom (which includes, meaningful freedom in areas of significance to one’s life) as being incompatible with autonomy – this includes, crucially, extreme cases where heteronomy seems a reality *de facto*, e.g. slavery or forced prostitution.

Second, Oshana highlights that a focus on psychological conditions or in subjective strategies should not be favoured when, by doing the latter, one would make little progress

---

<sup>432</sup> Ibid., 85

<sup>433</sup> Oshana, “Is Social-Relational Autonomy a Plausible Ideal?” 4.

<sup>434</sup> Ibid., 12.

in *enhancing* individuals' autonomy in conditions of oppression. Again, extreme cases like the ones mentioned above could be evoked as examples of the latter: being a slave or being forced into prostitution are conditions of heteronomy sustained mostly by *external* elements. Therefore, the degree to which one's autonomy in these conditions could be enhanced by appealing to a change in agents' psychologies seems minimal.

Having said this, however, I believe that Oshana's key insights (and, therefore, her substantive strategy) should be contained and applied only to cases of 'domination'. My intuition is that, when oppression does not amount to this extreme, then Oshana's points can be overcome. For example, when one is subjected to oppressive norms or stereotypes, one could both (i) have access to meaningful or significant freedom on areas of significance to one's life, and (ii) limit one's oppression meaningfully through subjective strategies. I develop and justify these points in the next and final chapter.

## Chapter 5: A Dual-track Approach to Autonomy

In light of the critical assessment of procedural and substantive models developed in Chapters 3 and 4 respectively, my aim in this Chapter is to provide an alternative that overcomes the difficulties raised. To do the latter, I defend a two-tracked strategy that opts for a substantive or procedural approach to autonomy depending on the *kind* of oppression in place. In a word, because oppression can take (qualitatively) different forms, I argue, a model of autonomy which intends to promote individuals' emancipation (i.e. to limit oppression) should not be fully committed to either a procedural strategy or to a substantive one alone. As I have anticipated, I propose that, in oppressive circumstances, agents' autonomies should be decided *procedurally* when they are 'subjected' but *substantively* when they are 'dominated'. A dual-tracked strategy, I argue, allows us to distinguish between oppressive states which *require* a particular form of self-management (i.e. subjection) and those that might benefit from self-management to some extent but do not ultimately need it (i.e. domination).

'Subjection' and 'domination' were defined in Chapter 2, by drawing from Foucault's analysis of power, let me start by summarizing how I have defined these categories and by explaining how they can be useful in the context of this discussion:

Even though Foucault does not explicitly deal with oppression, throughout my thesis I have argued that Foucault's framework is fruitful to distinguish qualitatively different forms of oppression. For example, some forms of oppression like slavery, colonialism, sexual exploitation can be likened to a notion of 'domination' inspired by Foucault's philosophy. Alternatively, other forms of oppression which do not typically involve violent or disruptive interventions on agents' developments (i.e. they affect agents' developments as a whole), and which are maintained by encouraging specific ways of self-relating, being, and living which are functional to the power configuration in place, can be thought as Foucauldian



‘governmental techniques’. As I have argued in previous chapters, some forms of race, class, and gender oppression fall within this second category. I use the term ‘subjection’ to refer to forms of oppression of the latter kind.

In Chapter 2, section III, I focused on Foucault’s attempt to distinguish power relations from other phenomena typically associated with power. In this context, I identified two constitutive features of power according to Foucault: (i) power is exercised over a “person who acts”;<sup>435</sup> (ii) power relations “structure” or “manage” fields of possible “actions, reactions, and results” but they do not close or limit them completely.<sup>436</sup> Crucially, the latter entails that power relations are reversible through the margin of freedom available to the agents within the power relation.<sup>437</sup> As I said in Chapter 2, power *influences* the way in which agents exercise their *freedom*.<sup>438</sup>

The characterization of power relations presented above makes it possible to distinguish power relations from other extreme situations where power has gone “too far”. For example, Foucault distinguishes power relations from violently “crushing” subjects and from cases where social power strategies have been “frozen”, e.g., through the use of force or violence. Also, power relations may have become so asymmetrical that those over whom power is exercised have their field of possible action extremely limited. That is, something other than power is going on in extreme cases where there is no (to put it in the terms presented earlier) “induction or seduction” of conducts but rather “bending and breaking” of spaces for possible action and/or of bodies.<sup>439</sup> I use ‘domination’ to refer to the latter cases and include within this category three phenomena identified by Foucault: ‘relationships

---

<sup>435</sup> Foucault, “The Subject and Power,” 340.

<sup>436</sup> Ibid.

<sup>437</sup> Ibid., 342; 346.

<sup>438</sup> Ibid., 341.

<sup>439</sup> Ibid., 340-1.

of violence’, macro-level ‘states of domination’, and extremely asymmetrical (and therefore “frozen”) micro-level power relations.<sup>440</sup>

By ‘subjection’ I refer to forms of oppression which can be likened to Foucauldian ‘governmental techniques’. Governmental techniques are forms of social power which direct individuals’ conducts by the “external” means associated with power (e.g. management of individuals’ fields of action, observation and regularisation of conducts, etc.) but also through ‘technologies of the self’.<sup>441</sup> Government, then, also relies on actions carried out by individuals on their own bodies, thoughts, and conducts with a view to shaping themselves in light of given (social) ends (e.g. a moral standard).<sup>442</sup> The “external” and “internal” dimensions of government are tightly connected because, as I explained in Chapter 2, governmental techniques govern individuals *en masse* (i.e. as populations) by encouraging that the individuality of each and every one of the members of a certain population submits “to a set of very specific patterns”.<sup>443</sup> Foucault claims that the ‘government of individualisation’, which marks Western societies after modernity, “ties the individual to himself and *submits him to others in this way*”.<sup>444</sup>

Techniques of government, however, are forms of social *power* and, as such, they structure but not exhaustively limit the spaces of possible actions of governed individuals. Importantly, the latter entails that the room for play available to those who are governed is enough to reverse or significantly limit government. Moreover, since government also relies on practices of self-direction, self-transformation, and self-constitution, the possibilities for

---

<sup>440</sup> See Ibid., 340 for a characterisation of ‘relationships of violence’; See Foucault, “*The Ethics of the Concern of the Self as a Practice of Freedom*,” 283 for a definition of (macro-level) ‘states of domination’; See Ibid., 299 for a description of (micro-level) ‘domination’.

<sup>441</sup> Foucault, *About the beginning of the hermenautics of the self*, 26.

<sup>442</sup> Ibid., 25.

<sup>443</sup> Foucault, “The Subject and Power,” 334.

<sup>444</sup> Ibid., 331.

significantly altering (and limiting) governmental techniques are within the reach of governed agents.

The reasons that justify the adoption of a two-tracked approach to deal with domination and subjection emerge from the analyses carried out in previous chapters:

First, a dual strategy is justified in light of the aims I set – namely, to develop a model of autonomy that allows us to identify the conditions of possibility for self-clarification and emancipation under oppression. In light of these aims, one concern has been to determine the emancipatory potential of different forms of self-assessment – crucially, it has been central to analyse how self-assessment could be made critical enough to enhance individuals' autonomies in contexts of oppression. For example, I have argued that in order to be able to detect frequent forms of oppression in contemporary societies (e.g. being affected by oppressive stereotypes) through reflective means, agents' self-assessments need to be informed by the right kind of *social* historical materials.

However, as it became clear in Chapter 4, there are also cases where self-clarification, no matter how critical or powerful, will *not* be sufficient to significantly limit oppressive heteronomy and its effects on relevant aspects of one's life. A key factor in these latter cases is that *external* limits prevent agents' lives from being shaped in light of self-governed decisions in any meaningful way. I claim that such extreme cases are those which I have called 'domination'.

Two crucial features justify, then, the adoption of a non-procedural strategy when oppression amounts to domination, namely that (i) when agents are dominated, there is no room for *meaningfully* shaping one's own life; and (ii) under domination, the key constraints that individuals face are typically due to material or external conditions outside their control, and not (merely) to reflexive attitudes or ways of conducting themselves.

For the sake of coherence, the first issue that I would like to address is whether making autonomy categorically incompatible with domination contradicts my previous claim on the

importance of avoiding a slide towards ‘orthonomy,’ i.e. towards an account of government by “right rules”. (To recall, I have committed to providing an account of *self*-government, where those “self-given” rules are not simply defined by ideals of what is, say, “objectively good” for individuals). I argue that my proposal does *not* conflate autonomy and orthonomy, and this for the following reasons:

First, as I have suggested above, the main reason for deciding autonomy substantively in cases of domination is that, under states of domination, individuals lack the kind of *control* over relevant aspects of their life that is *needed* for any form of *self*-government. Without this control, as I have argued when discussing Oshana’s position in Chapter 4, one lacks the means to significantly shape one’s life in light of one’s values and commitments. In view of the latter, I argue, when domination is the case the most urgent thing to do in order to secure *self*-government, is to advocate for change of those external circumstances that make agents heteronomous *de facto*. This could imply, for example, advocating for change of the material or legal conditions that secure domination.

These points will become clearer below, but let me for now briefly illustrate what I mean by recalling an example from Chapters 3 and 4:

Consider the case of Kaew, who was forced into prostitution but eventually “accepted” or “endorsed” her situation. As I noted in Chapter 4, the fact that Kaew’s life options are so severely restricted (and, indeed, that they were restricted through the use of illegitimate interferences) should be the decisive factor to determine Kaew’s heteronomy, regardless of the fact that Kaew seems (at some point of her life) to “accept” her situation. The reason for this is that it is difficult to see *any* choice or endorsement as a meaningful one when, for example, should Kaew’s motivations, desires or preferences change in the *slightest* way, she would be most likely exposed to violence or to other forms of direct coercion that would be obviously illegitimate.

Moreover, in cases like Kaew's, the likelihood of facing violence or coercion in the event of any change of preferences should be a reason to at least question the legitimacy of the endorsement of one's situation. What I mean is that, in cases like these, the constant threat of coercion and violence casts a shadow over the "voluntary" character of a person's choices. Similarly, even conceding that Kaew's endorsement of her situation could make her domination "easier" for her exploiters, it is not her endorsement that ultimately *causes or secures* her domination: coercion and violence (or the threat thereof) do.

And why are other cases of oppression that fall short of domination different? As I have suggested before, when by oppression we mean 'subjection', social power relations are significantly open-ended. Crucially, the key difference is that when individuals' oppression does not amount to domination, oppression is (at least to a significant extent) reversible through means that are within the reach of those oppressed, i.e. self-emancipatory practices are possible. As a result, individuals who are subjected but not dominated retain at least the necessary control for self-government to be meaningful – i.e. lives could be significantly shaped or re-shaped in light of different preferences, values, actions, and self-relations.

As I have mentioned several times, saying that subjection can be "reversed" does not imply that one could get rid of all power relations completely. As I put it in Chapter 2, it is impossible not to be governed *at all*.<sup>445</sup> Still, even if a life completely unaffected by power is impossible, it is possible to become *less* governed (i.e. to be more autonomous) vis-à-vis configurations that one finds particularly problematic or undesirable. For example, it is possible to realize that certain ways of being make me functional to social ends or structures that I reject and to alter the way in which I live and relate to myself (and to others) accordingly.

---

<sup>445</sup> Foucault, *What is Critique?*, 28.

Still, we should avoid downplaying the significance and the potential that limiting subjection can have: while limiting subjection is certainly not a way to abolish all forms of oppression completely, I argue that limiting subjection can significantly contribute to the external and structural change that theorists of oppression and (some) theorists of autonomy rightfully see as necessary to limit oppression. With this I mean several things:

First, limiting subjection should not be understood exclusively in an internalist fashion. For example, fighting against forms of subjection like gender oppression does not merely mean that agents should break their “inner selves” free from certain problematic social “truths” (e.g. by changing their preferences or reflexive attitudes). Claiming that only “internal” de-subjection is relevant would make us fall back on the one-sided view of oppression that I problematized in Chapter 4, when analysing ‘internalized oppression’.

Indeed, in order for de-subjection to be more than, say, the abandoning of a social stereotype for the adoption of a different ready-made one, it is both necessary that new roles become available and that available ones are “de-normalized”. For this to happen, however, change cannot be restricted to one occurring in individuals’ “inner” private worlds, insofar as social roles and identity categories are intersubjectively defined and regulated. Consequently, pushing back on oppression as subjection needs to involve an external or public dimension – e.g. it is also necessary to challenge the way in which available social roles are *enacted* and to promote the creation and enactment of new roles or ways of life.<sup>446</sup> For example, as Rosa Luz’ performance (presented in Chapter 4) shows, limiting her heteronomy through subjective means is not exclusively a “private” matter – i.e. it is not just a matter of, say, seeing herself differently but also of *acting* differently. Furthermore, limiting her

---

<sup>446</sup> Foucault’s classic example on this regard is the creation of the category of ‘gay’ which enabled new “ways of life”. Foucault claims: “A way of life can be shared among individuals of different age, status, and social activity. It can yield intense relations not resembling those that are institutionalized. It seems to me that a way of life can yield a culture and an ethics. To be ‘gay,’ I think, is not to identify with the psychological traits and the visible masks of the homosexual but to try to define and develop a way of life.” (Foucault, “Friendship as a Way of Life,” in *Ethics, Subjectivity, and Truth, Vol 1 of The Essential Works of Michel Foucault 1954-1984*, ed. Paul Rabinow (New York: New York Press, 1997), 138).

heteronomy is not exclusively a “personal” matter – i.e. she also tries to shake the social stereotypes and categories which condition and limit her (and others’ like her) appearance in public spaces.

Second, it is important to remember that, as I explained in Chapter 2, pushing back on ‘subjection’ (i.e. on governmental techniques) is also indirectly pushing back on domination. That is, there is an interest in encouraging a counter-governmental attitude that is not merely that of fighting normalisation or regularisation for the sake of, say, multiplying available ways of being. Furthermore, pushing against forms of subjection is also a form of avoiding that dominant ways of being become so “natural” or static that the power imbalances maintained by these forms of being end up ossified. In Foucault’s words: “[t]he analysis of [techniques of government] is necessary because it is very often through such techniques that states of domination are established and maintained.”<sup>447</sup>

Therefore, the project of pushing back against forms of subjection should not be reduced to a (collective or individual) fight against “abstractions”. Indeed, when Foucault himself reflects on the struggles against the ‘government of individualisation’ distinctive of his time (he writes this piece in the early 1980s and refers to feminist movements and to anti-psychiatric movements), he sees the latter both as a fight against the imposition of ways of being *and* against the different forms of violence associated with or enabled by these ways of being. Foucault claims:

... all these present struggles revolve around the question: Who are we? They are a refusal of these abstractions, of economic and ideological state violence which ignore who we are individually, and also a refusal of a scientific or administrative inquisition which determines who one is.<sup>448</sup>

---

<sup>447</sup> Foucault, “The Ethics of the Concern of the Self as a Practice of Freedom,” 299.

<sup>448</sup> Foucault, “The Subject and Power,” 331.

Pushing back against the government of individualisation, then, could also be seen as a way of unveiling that many of the abuses which derived from the practical implementation of these abstractions (e.g. psychiatric violence, marital violence, economic deprivation, etc.) were *unjustified*. Moreover, my point is also that by showing the artificiality of these “scientific” truths about people, one is also showing the arbitrary and unjustified nature of many of the regular administrative, economic, and institutional mechanisms that keep individuals “tied” to these truths.

These points could be illustrated (and made more concrete) by turning to Bartky’s remarks on (what she calls) ‘psychological oppression’ which, as I explained in previous chapters, shares some crucial features with ‘subjection’ (as I have defined it). Bartky claims:

Every mode of oppression within the system [psychological, political, and economic] has its own part to play, but each serves to support and to maintain the others. Thus, for example, the assault on the self-esteem of white women and of black persons of both sexes prepares us for the historic role that a disproportionate number of us are destined to play within the process of production: that of a cheap or reserve labor supply.<sup>449</sup>

The point that I want to make in light of Bartky’s claim is that seeing ourselves as subjects of different sorts (e.g. seeing ourselves as subjects of different “moral worth”) makes certain forms of life that may otherwise appear unbearable, unjustified or unappealing for those

---

<sup>449</sup> Bartky, “On Psychological Oppression,” 32; emphasis added. Admittedly, the conclusions that Bartky draws from this observation could differ, in some degree, from mine. Namely, Bartky would not consider that being governed in light of a “truth” about who we are (or, in her terms, being ‘psychologically oppressed’) could be significantly reverted without changing the material conditions that sustain psychological oppression. She claims later on the same page: “Because of the interlocking character of the modes of oppression, I think it highly unlikely that any form of oppression will disappear entirely until the system of oppression as a whole is overthrown.” (Ibid.) As it emerges from the ongoing discussion, I am sympathetic to this kind of observation and generally in agreement with Bartky. As I have signalled before my emphasis is on the possibility of significantly *limiting* one’s heteronomy, which (I argue) *can* be accomplished by subjective means when oppression is likened to ‘subjection’ but falls short of ‘domination.’ It should be clear then that I am not committed to a view in which merely by subjective means, *any* form of oppression will *disappear entirely*.



concerned appear “acceptable” or, even, “natural”. As a result of the latter, the economic or legal arrangements that negatively affect those falling within a certain “social kind” are likely to find less social resistance. For example, being socialised as a ‘woman’ may diminish individuals’ resistance to the implementation of patriarchal laws (e.g. severe bans on abortion) or increase one’s tolerance to some forms of violence.<sup>450</sup> Race stereotypes may serve similar effects by, for instance, making economic, legal, and physical violence appear acceptable or justified.

Third, opposing domination and limiting subjection are related in another important way. As Foucault explained, ‘states of domination’, like colonialism, require liberation but, once liberation happens, it is *also* crucial that those formerly dominated work on themselves and shape this new space of possibilities in a way that enables a new (political, social, ethical) existence.<sup>451</sup> That is, once ‘liberation’ from severe external constraints is accomplished, it is necessary that ways of being and living actively challenge the (now not so radically determined) spaces of possibilities. Foucault claims:

...liberation is not in itself sufficient to define the practices of freedom that will still be needed if this people, this society, and these individuals are to be able to define admissible and acceptable forms of existence or political society.<sup>452</sup>

So, once those crucial external barriers or conditions that enable domination are reversed, a task of self-clarification will be necessary and apt.<sup>453</sup>

---

<sup>450</sup> See my example in Chapter 3 on marital rape.

<sup>451</sup> Foucault resisted the idea that ‘liberation’ could send subjugated people back to an “original” state (i.e. a human nature) from which these people had been alienated due to historical or material conditions. (Foucault, “The Ethics of the Concern for Self as a Practice of Freedom,” 282.)

<sup>452</sup> *Ibid.*, 282-3.

<sup>453</sup> See McWhorter, “Post-Liberation feminism” for the implications of this point for feminists’ struggles. See Saidiya Hartman, *Scenes of Subjection: Terror, Slavery, and Self-making in Nineteenth-Century America* (New York: Oxford University Press, 1997) for an analysis of the economic, social, and legal effects of the processes of “humanisation” and “individualisation” of former slaves after the abolition of slavery in the USA.

In a word, limiting subjection is not just a matter of being governed *otherwise*, full stop. Instead, setting limits to the ways in which we are governed should also mean, as Foucault suggests, seeking to play power with “as little domination as possible”.<sup>454</sup> This implies, for example, that identity categories or social roles should be (both individually and collectively) challenged with an eye on the part that they play (or could play) in the implementation of, for instance, forms of material exploitation or of legal neglect of some populations. It should be clear then that the fact that my proposal outlines two different tracks should not be read as the pursue of two different aims. Rather, my dual-tracked strategy delivers the promise made at the beginning of this thesis, namely developing a finer-grained analysis of (what contemporary critics call) ‘oppression’ and using that analysis to shape a model of autonomy that is strong enough to limit oppression in its various forms.

All in all, I believe that the tipping point which differentiates domination from subjection could help advance the contemporary debate about theories of autonomy beyond its current impasse between opposing, procedural and substantive, factions. The idea is that limiting heteronomy as oppression (the end behind this two-tracked model) requires a substantive strategy when domination is the case, but should imply a (revised) procedural one (i.e. should be decided by the agent whose autonomy is at stake) when one is dealing with subjection.

Let me analyse each track of my model in turn.

### **I. Substantive Track – Oppression as ‘Domination’:**

As I explained above, when oppression could be likened to ‘domination’, its perpetuation does not ultimately rely on specific forms of agency. What I mean is that the

---

<sup>454</sup> Foucault, “The Ethics of the Concern for Self as a Practice of Freedom,” 298.

ways in which agents relate to themselves and direct their conducts is not the key factor keeping them in a position of subordination. For example, when one's possibilities for action are severely restricted by a racist or sexist law that is strictly applied, when political dissidence is suppressed through a dictatorship, or when an individual is controlled through force and violence, one is typically confronted with domination. In these extreme cases, I argue, the settled ways of being and acting of the oppressed are not the key factors that need to be addressed to limit oppression.

To be sure, the conducts of dominated individuals need not be completely irrelevant. Indeed, someone could rightfully argue that even in very extreme cases (e.g. a totalitarian regime seeking to eliminate political dissidence), non-conforming agency is seen as a danger (e.g. a totalitarian regime is very aware of the risk posed by a dissident's actions when she is incarcerated). However, my point is that, in these extreme cases, non-conforming or dissident agencies are "managed" very differently than they are in cases of subjection, namely in the former case they are not produced and directed but, rather, externally repressed, silenced or immobilized. Indeed, even in cases of domination where agents are not completely "crushed" (we could even imagine, for the sake of the argument, agents who are dominated but who retain the "freedom of mind" that they would need to critically assess their conditions), *external* possibilities for action are blocked in a way that does not allow for significantly different forms of living and acting. In a word, it is the *fact* that one's margin of possible action is extremely reduced which sustains domination and not the use one may make of any margin of freedom left.

Take the example of the European marital structure of the 18<sup>th</sup> and 19<sup>th</sup> centuries which, as explained in Chapter 2, Foucault deems a case of domination. As Foucault notes, even when women could adopt different strategies to deal with their situations, the fact remains that very restrictive (legal and economic) barriers were still there at the end of the

day. Therefore, it is still true that private or individual strategies were not enough to significantly *reverse* the state of domination.

Admittedly, someone could worry that by automatically ruling out the possibility of dominated people to be autonomous, I am dismissing heroic forms of agency which may be rare but not impossible. Think for example of those who resist extreme forms of abuse, e.g. some slaves or political prisoners, by managing to “stay true” to themselves and their values in a way that *does have* some effect on their lives. For instance, they may avoid going mad or retain a minimal sense of self-esteem. Are we missing something crucial if, for the purposes of deciding autonomy under oppressive circumstances, we ignore subjective strategies and coping mechanisms like these?

In reply, let me specify that I grant that even though domination cannot be reversed by appealing exclusively to subjective strategies, the annihilating effects of some extreme forms of domination may nevertheless be resisted in different degrees. What I mean by this is that an agent may manage to retain a repertoire of capacities and personal features that constitute the conditions of possibility for agency (autonomous or not) even when there is not sufficient margin of freedom to exercise that agency meaningfully. This on its own, however, does not guarantee autonomy – one may feel inclined to say that such agents are in a better position or better suited for autonomy should liberation come, but the fact that one’s freedom is *severely* restricted means that they *currently*, in their state of domination, lack it.

I illustrate my points by considering a case of political imprisonment. The memoirs of Rosencof and Fernández Huidobro, prisoners in Uruguay from 1972-1985 (before and during the Uruguayan dictatorship), allow me to clarify why and how a procedural approach is inadequate to capture the key autonomy-undermining features in extreme oppressive cases. The case of political imprisonment that I present below qualifies as an instance of (what I call) ‘domination’ for two reasons:

First, on the macro-level, the political and social scenario in which imprisonment occurred could be seen as a ‘state of domination’ in the Foucauldian sense: a military class “blocked” social power relations. Any individual or collective attempt (no matter how small) to “subvert” the military regime entailed the risk of imprisonment, torture, and forced disappearance. Therefore, even when not imprisoned, political dissidents (and anyone considered to be one) lacked control over the most basic aspects of their lives in a way that could not be significantly limited through the actions available to them.

Second, when imprisoned, political dissidents were subjected to sheer violence: political imprisonment was not aimed at “taming” the subjectivity of dissidents or at “normalising” them. Indeed, the kind of violence which characterised these cases was meant to “break” bodies and subjectivities, not to “re-shape” them in a more “convenient” way.

Let me now introduce my example:

For twelve years Rosencof and Fernández Huidobro lived under the most appalling circumstances: they had to survive in inhuman material conditions and were victims of torture and other forms of physical and psychological abuse. The (sadly) common practice of “disappearing” political dissidents was not an option for the regime in their case. (This was so for a range of reasons, most notably because they had been imprisoned before the beginning of the official dictatorship and appeared on official registers. Rosencof and Fernández Huidobro were also “high-profile” prisoners –former leaders of the *Movimiento de Liberación Nacional-Tupamaros*—<sup>455</sup> and, as a result, they became useful hostages for the dictatorship: should their comrades outside rebel, they would suffer the consequences.) According to the claims of high-rank officers at the time, since they could not kill these prisoners, orders were given to “drive them mad”.<sup>456</sup>

---

<sup>455</sup> Spanish for “National Liberation Movement-Tupamaros”.

<sup>456</sup> Eleuterio Fernández Huidobro and Mauricio Rosencof., *Memorias del Calabozo* (Navarra: Ed. Txalaparta, 1993), 21

Somehow aware of this situation, Rosencof and Fernández Huidobro made it their top priority to “show how a human being, [...], can resist appalling cruelty without becoming a beast or a plant.”<sup>457</sup> During the time of their captivity, these two men managed to retain not only some minimal conditions for agency (e.g. the necessary reflective skills to think for themselves) but also the perspectives they called their own. Indeed, fighting against the annihilation of their values and *personalities* is presented by the authors as a sign of resistance.<sup>458</sup>

This exceptional case may seem to work as a counterexample of my dual-track solution: their imprisonment does not depend on the decisions they make or on how to conduct themselves, the margin for carrying out self-governed actions is extremely reduced both in alternatives and in scope (i.e. the margin for freedom is not only minimal but also limited to dimensions of their lives that, under different conditions, they would not consider relevant),<sup>459</sup> and *yet* one is inclined to say that psychological resistance *does* make a difference.

The question we should ask is, nonetheless, *what* is exactly limited by subjective means in this case? Do they become less heteronomous vis-à-vis their domination by avoiding being “crushed” by their oppressors? I argue that the fact that external and structural circumstances are not significantly altered (i.e. they continue to be dominated) should remain paramount here. As relevant and admirable as avoiding being “crushed” may be, it is not merely a capacity for functioning agency that is necessary to enhance one’s autonomy under oppression. Additionally, it is necessary that agents retain a meaningful space to exercise that agency and shape their lives in light of their values. In a word, as I argued in Chapter 4

---

<sup>457</sup> Ibid.; my translation.

<sup>458</sup> For instance, they encouraged each other to imagine a sign inside their cells reading “Acá también se lucha [In here, we are also fighting]”. They report that being able to self-identify as “fighters” or as other fellow-fighters’ “comrades”, even in their prison cells, gave them a sense of dignity that they consider crucial for survival. (Ibid., 74)

<sup>459</sup> For example, they explain how, because of their limited access to toilets, developing strategies to urinate in their cells became a very central part of their everyday lives and was necessary to avoid making the state of their cells even more unliveable. Retrospectively, they find it difficult to believe how much mental energy and planning such a basic human need required. (Ibid., 82)

following Oshana, autonomous agency needs minimal control over relevant aspects of one's life. Therefore, resisting annihilation is not equivalent to limiting (oppressive) heteronomy.

A crucial (subjective) difference between resisting violence or annihilation under domination and limiting oppression should be noted. While, as I have argued in previous chapters, limiting oppression necessarily involves challenging (or, at least, having had opportunities for challenging) sustained values and ways of being, the criterion of "resistance" presented above is much less demanding. We need not engage in subtle evaluations (like asking ourselves what "kind of self" is preserved) to assess resistance under domination. For example, an individual could resist annihilation by holding to a self that is the result of a process of subjection that she has not questioned. In other words, while for resistance it might be sufficient to hold to the values that were there before (if applicable), for autonomy one's long-term (i.e. stable, sustained) commitments and character features need to be problematized too.

Finally, let me point out that my proposal can accommodate the experiences of those living under less-extreme (and, admittedly, more common) forms of domination. Consider those campaigning for external (e.g. political, economic) change: it is the fact that agents "hit a wall" when trying to shape their lives differently or to orient their actions differently that makes them realize the need to "liberate" themselves from certain external constraints. It is also generally true, I would say, that considering these people heteronomous is compatible with their *own* interpretation of their situation: a quest for external change typically implies the acknowledgement (and the discomfort, worry or rage) that one's life is limited by unacceptable constraints. Briefly: saying that one is not autonomous under domination need not disregard individuals' experience of their situation.

Admittedly, a more challenging scenario could be considered. Indeed, when theorists of autonomy worry about the consequences of externally deeming an agent heteronomous (i.e. of doing so regardless of *her* perspective on the matter) they have in mind cases in which

an observer's judgement conflicts with the agent's: someone deems heteronomous an agent who judges her own situation acceptable or autonomous. In this particular case, my strategy could be seen as disregarding the perspective of agents who might reflectively endorse their domination.

I believe that even if we conceded that the scenario above is possible, my proposal could not result in any serious damage to the agent given the scope and aims of my model – namely, it would not be on the grounds of deeming an agent autonomous or heteronomous that basic rights would be adjudicated or denied. Moreover, the (understandable) worry that agents' rights or perspectives should not be disregarded is actually taken more seriously, I believe, if one acknowledges the need to dismantle particularly damaging circumstances (i.e. domination) that *de facto* deny individuals basic rights and/or meaningful options.

Before concluding this section, an important qualification is needed. It is certainly possible to imagine oppressed individuals whose lives are completely dominated – in all relevant aspects, they are in the “hands of others,” such that their own actions cannot significantly reverse this. The life of some slaves might be an example of this. It is also possible to imagine oppressed individuals who are free from all domination vis-à-vis relevant aspects of their lives: someone subjected to oppressive gender norms in a society where possibilities to live otherwise are imaginable and feasible within the space of possibilities available to them would be such an example. However, some individuals' lives might combine forms of subjection *and* forms of domination.

Consider for example the case of women trying to deal with unwanted pregnancies in countries where abortion (or presumption thereof) is severely penalised (e.g. El Salvador). It is possible to think that under these circumstances some women may retain sufficient control over relevant aspects of their lives and their sexualities to push back against some dimensions of gender oppression. For example, they might critically assess social gender norms which connect their self-worth to motherhood and shape their lives in light of this critical



assessment. Still, women might confront important circumstances with significant effects over relevant aspects of their lives (e.g. an unwanted pregnancy) in which limiting the influence of an oppressive structure on themselves would be beyond any subjective strategies they might adopt. Cases like these would reaffirm what I have claimed throughout my thesis, namely that autonomy is *not* an all-or-nothing matter. The fact that one may suffer domination vis-à-vis some aspects of one's life and subjection vis-à-vis others, is fully compatible with this.

## II. Procedural Track – Oppression as ‘Subjection’

I argue that a procedural approach to autonomy may be suitable to tackle oppression as subjection. Under these circumstances, as stated above, the key problem is that individuals self-relate, self-fashion, and make use of the margin of freedom available to them in a way that helps sustain their oppression. As a result, a form of critical self-assessment that guarantees sufficient conditions for a subject to be able to revise the conditions of her subjection, could have promising emancipatory effects. As I have argued in previous chapters, however, not all forms of self-clarification and self-assessment are effective to tackle subjection.

In this section, I reflect on which conditions need to be met by a procedural model of autonomy to ensure that agents are able to *critically* assess themselves even when subjected. To determine the latter, I illustrate how Christman's model could be strengthened in light of my aims.

As I have argued in previous chapters, Christman's historical model is a promising alternative to deal with forms of oppression that affect an agent's development over time. Historical models are right to point out that one's history is key information to carry out self-assessment. For example, as Christman acknowledges, when it comes to see if one deeply

rejects dispositions or values, it does make a difference to test one's assessment of these dispositions or values in light of how and why these contents originated. However, as I argued in Chapter 3, Christman's historical model does not take the critical stance vis-à-vis one's history or one's stable ways of being that would be necessary to limit subjection. I have justified this claim as follows:

First, I have argued that, in Christman's account, one's personal history serves more explanatory or justificatory than critical purposes. For example, in Christman's most developed account, one's development typically allows to grasp those organizing values, affective reactions, and ways of thinking that, because of their stability over time, should be taken as reference when assessing preferences or dispositions.<sup>460</sup> This, together with the fact that historical models leave underspecified what kind(s) of history (personal, social, political) is necessary for autonomy, leaves Christman's historical conditions under-prepared to deal with forms of oppression which affect one's general development as a subject.

Second, I have argued that Christman's model relies a great deal on agents' experiences of deep alienation without sufficiently ensuring that the conditions of possibility for experiencing alienation are met. The latter is particularly pressing in light of the possibility of subjection. That is, it should be acknowledged that experiencing alienation vis-à-vis practical identities which may be (socially) presented and accepted as truths on who one is, and which may have shaped one's development through and through, may not be a straightforward matter. For example, I have argued that full (affective and cognitive) "acceptance" of a social role is compatible with subjection and that, therefore, the absence of experiences of alienation should not be taken as a sign of autonomy without making sure that agents have the (affective and cognitive) resources for *testing* their acceptance. Furthermore, I argued that to experience discomforts or strong negative experiences *as*

---

<sup>460</sup> Christman, *The politics of Persons*, 153-4.

alienation, agents need a minimal understanding of the social roles that they enact and of the social function that these roles play.

Ensuring the conditions of possibility for experiencing alienation is all the more important in light of Christman's revised model. As I explained in Chapter 3, according to Christman's revised account, the self-assessment necessary for autonomy continues to be a reflective one (i.e. an autonomous agent should have the "disposition to reflectively accept one's practical identity").<sup>461</sup> However, for the purposes of confirming this reflective validation, Christman argues, one can typically "trust" the absence of experiences of alienation.<sup>462</sup> Therefore, in Christman's revised account, the agent's recurrent positive experiences vis-à-vis her actions in light of her identity are taken as good evidence of the existence of *reasons* for living according to a certain practical identity. This, I have argued, could set the bar too low for autonomy if agents have not had opportunities for testing recurrent experiences and settled commitments.

In light of the abovementioned features, I claimed that the critical potential of Christman's historical model is limited to deal with subjection. The above are, I argue, the aspects that need to be strengthened if a procedural model is to grant agents the reflective and affective materials they need to critically assess themselves under subjection.

For the sake of clarity, let me illustrate Christman's points, my worries, and my suggestions through an example from literature: *The Remains of the Day* by Kazuo Ishiguro.<sup>463</sup> Ishiguro's main character, Stevens, represents for some autonomy theorists a clear case of someone whose life may seem tragic or painful to *others* but who should, nonetheless, count as autonomous. Available readings of this case focus on showing that Stevens satisfies all the

---

<sup>461</sup> Christman, *Decentered Social Selves*, 53.

<sup>462</sup> *Ibid.*, 54.

<sup>463</sup> Kazuo Ishiguro, *The Remains of the Day* (London: Faber and Faber, 1989).

necessary (reflective and reflexive) procedural conditions for autonomy and suggest that, deeming Stevens heteronomous, would necessarily imply adopting content-laden standards regarding which lifestyles *should* count as autonomous.<sup>464</sup>

I suggest an alternative reading of this case: Stevens qualifies as (procedurally) autonomous only because his perspective does not sufficiently consider the *social* significance and meaning of his role. As a result, Stevens lacks the resources to articulate the discomfort or suffering that he does experience into the fully-fledged experiences of alienation that Christman considers a marker of heteronomy. Gaining cognitive and affective resources to experience alienation includes, as I argued in Chapter 3, both becoming capable of making sense of unarticulated experiences of discomfort, and becoming capable of seeing other more obvious negative experiences (e.g. regret or guilt) that may have been deemed personal *as* experiences of alienation.

Stevens, the main character of the novel, tells his story in 1950s England, when he is the butler in (what used to be) a traditional English country estate called “Darlington Hall”, now the property of a wealthy American (Mr. Farraday). Stevens finds himself confronted with both the fact that there does not seem to be a place for a traditional English butler in post-war Britain and with his incapacity to live up to the standards of this role (which are also *his own* standards) because of his age. A letter from Miss Kenton – an old friend and a former housekeeper in Darlington Hall – motivates a road trip to visit her. During this trip, Stevens goes through the memories of his life in Darlington Hall. Stevens’ past is recalled vividly and with great detail and one can see why Stevens would pass Christman’s procedural test of autonomy (in the two versions presented above) with flying colours. Let me mention some crucial aspects of Stevens’ history that justify this claim:

---

<sup>464</sup> Christman considers this example briefly in Christman, *Decentred social selves*, 53-4. A thorough analysis of the novel and of its main character can also be found in Kwame Anthony Appiah, *The Ethics of Identity* (New Jersey: Princeton University Press, 2005.) Appiah’s analysis differs significantly from mine. I comment briefly on Appiah’s analysis below.

First, Stevens generally appears ‘non-alienated’ from his values in the sense described by Christman: Stevens’ description of his past life during the golden years of Darlington Hall gives no evidence of the acute alienation that for Christman is a sign of heteronomy. Moreover, this non-alienation follows a self-assessment that happens in full awareness and even acceptance of his personal history (e.g. Stevens takes pride in having followed his father’s steps in becoming a butler).<sup>465</sup> Furthermore, nothing in the genesis of Stevens’ values suggests that procedural independence conditions might have been violated: Stevens’ life was not subjected to illegitimate interferences like coercion or manipulation. Indeed, in the majority of the novel, his life “smoothly” unfolds as socially expected *for a person like him*.

Second, throughout the novel Stevens provides evidence of the generally positive reflexive relation to his identity during his time in service. He repeatedly remembers the *great* butler that he was and makes remarks on the kind of dignity attached to his being a butler and to fulfilling his professional duties as well as he did. Therefore, when in service (and to some extent throughout his trip as well, I would add) Stevens *reflexively* affirms himself: he orients his actions from a specific social position which gives value to those actions and the latter confer value back to his role as a butler.

Indeed, Christman’s own analysis of the novel confirms the above. Christman claims that Stevens “remains autonomous” in most of the novel.<sup>466</sup> The key reason for this is that, the isolated feelings of “conflict” that Stevens might have experienced during his time in service, did *not* succeed in “short-circuiting his motivational/reflective feedback loop”.<sup>467</sup> Christman then concludes that, during the main period of the novel, “[Stevens] sense of

---

<sup>465</sup> Indeed, he seems to overall *glorify* his past to the point that some even attribute to Stevens a “post-imperial melancholy”. See Graham MacPhee, *Postwar British Literature and Postcolonial Studies* (Edinburgh: Edinburgh University Press, 2011), 144.

<sup>466</sup> Christman, *Decentred Social Selves*, 53-4.

<sup>467</sup> *Ibid.*, 53.

himself as leading an honorable life of obedience continues to affirm itself in functionally effective ways”.<sup>468</sup>

However, as I suggested above, it is hard to go through the novel without the sense that there is something tragic about Stevens’ life. Indeed, procedural theorists like Appiah and Christman attempt to account for our “intuitive” reactions to Stevens’ story in different ways, both of them faithful to the procedural commitment of giving agents the final say on their autonomy:

Appiah believes that even if Stevens’ life is not one that many would approve of or choose, Stevens’ is an example of the “moral power of individuality”.<sup>469</sup> Stevens works hard on his self-development according to a chosen life plan and *is* autonomous regardless of *our* disapproval of his life.<sup>470</sup>

Christman suggests that what strikes as tragic about Stevens’ life can be spelled out by appealing to Stevens’ *own* point of view and assessment of his life, which seems to change after his road trip. Christman notes that, towards the end of the novel, Stevens “is induced to reflect on that life and feels the tragedy of it [...] At that point, there is a crisis, and his reflections on himself and his life do cause him great misery.”<sup>471</sup> I want to emphasize that Christman does not suggest that Stevens shifts from a non-reflective acceptance of his life to a reflective rejection of it: First, as I explained in Chapter 3, reflexive self-assessment in Christman’s revised model typically evidences a disposition to reflectively accept one’s identity. Second, what Christman suggests is that Stevens rejects his life *retrospectively* but that “during his time in service [Stevens] was autonomous as an obedient servant who does not reflect in any way more deeply than level 2”.<sup>472</sup> (To recall: Level-2 reflection considers what

---

<sup>468</sup> Ibid., 53-4.

<sup>469</sup> Appiah, *The Ethics of Identity*, 13

<sup>470</sup> Ibid., 12

<sup>471</sup> Christman, *Decentred Social Selves*, 54.

<sup>472</sup> Ibid.

one has reason to do “[g]iven the kind of person I *am* and the commitments I have”, i.e. it typically *uses* a value framework *given* by one’s self-conception in a social context).<sup>473</sup>

As I see it, Stevens’ autonomy vis-à-vis his role as a servant throughout the novel is not so clear. However, my point is also that we do not need to make a move incompatible with a procedural view to show why Stevens’ self-assessment might be inconclusive (i.e. we do not need to adopt a substantive view to signal a possible tension). Rather, I argue that an unarticulated discomfort was *already* there during some of the past events that Stevens recalls. Stevens’ impossibility to articulate his own experiences of discomfort, I argue, is due to the kind of analysis of his life he engages with in most of the novel, which is too “narrow”.

Stevens’ reflective self-assessment, as Christman correctly points out, hardly ever goes beyond Level 2. As a result, Stevens is frequently able to find good enough reasons for his actions *given* his role and *given* his commitments, but he never succeeds in *challenging* them, not even when these commitments start to cause some discomfort. Moreover, part of the issue is that the instances of discomfort that he does experience are quickly dismissed or interpreted as evidence of his incapacity to live up to the standards of his role. In other words, at all times Stevens privileges maintaining “functional effectiveness”<sup>474</sup> (to put it in Christman’s terms) over challenging his commitments. What I mean is that “staying the same” is simply the easiest thing to do socially (to keep his life functioning and not to lose access to forms of social praise) and probably psychologically (admittedly, reshaping one’s values does not come without psychological cost, I say more about this below). Let me illustrate these points through a few examples from the novel:

At one point of the novel Stevens acknowledges that he prevented his relationship with Miss Kenton from becoming more personal by a series of “small decisions” that were, at the time of the events, hard to explain to himself. For instance, when Stevens recalls an

---

<sup>473</sup> Ibid., 51.

<sup>474</sup> Ibid., 54.

abrupt decision to end some professional evening meetings with Miss Kenton, he shows more than retrospective regret: he also reports some difficulty in making sense of this decision and its “full implications” at the time.<sup>475</sup> Additionally, Stevens describes a situation where his incapacity to break professional protocols to show empathy with Miss Kenton (who had lost a close relative) certainly caused him “strange feelings”.<sup>476</sup>

For example, one element is never considered during Stevens’ self-assessments (and by ‘never’ I mean: neither at the time of the events nor at the moment when Stevens tells his story): that pursuing a personal relationship with Miss Kenton was somehow incompatible with their role as live-in servants and, therefore, pursuing a personal relationship would have forced them to inhabit different social roles, to reconfigure their values, and to potentially lose some of the “dignity” that was available to them.

Significantly, Stevens’ over-engineered<sup>477</sup> discourse seems difficult to maintain when he starts gaining distance (literal and metaphorical) from the place and the role that he has always occupied. For example, at one point of his trip Stevens is even taken for *someone else* and this episode proves “taxing” and “discomforting”.<sup>478</sup> Stevens runs out of petrol and is sheltered by a couple of villagers who take him for a Lord (Stevens is driving an expensive car that he borrowed from Mr. Farraday, he is wearing a suit that had been left behind by a former guest of Darlington Hall, and speaks impeccable English). Over dinner, a discussion about “dignity” comes up and, even if Stevens sees his opinions respected, he is also challenged by a villager who defends his right (i.e. the right of a person like him) to have an

---

<sup>475</sup> Ishiguro, *The Remains of the Day*, 184.

<sup>476</sup> *Ibid.*, 186. More examples of these discomforts, perplexities, or even “quiet desperations” can be found throughout the novel vis-à-vis central aspects of Stevens’ life (e.g. his relation to his different masters, with his father, among others).

<sup>477</sup> I borrow this term from MacPhee’s analysis of Stevens’ narration: “Stevens’ narration roams obsessively over his field of memory, avoiding what is difficult or painful and displacing unacknowledged anxieties into elaborate descriptions of trivial details or over-engineered interpretations of incidental event and bogus points of principle.” (MacPhee, *Postwar British Literature and Postcolonial Studies*, 144.)

<sup>478</sup> Ishiguro, *The Remains of the Day*, 189-90.



opinion *even* in front of a Lord.<sup>479</sup> Even though Stevens initially rejects the more democratic ideals of the villager, later in the novel he comes back to the subject of “dignity” and reflects:

All those years I served him [Lord Darlington], I trusted I was doing something worthwhile. I can’t even say I made my own mistakes. Really – one has to ask oneself – what dignity is there in that?<sup>480</sup>

At this point, then, we can clearly see a full breakdown of his value framework, i.e. alienation in the sense of Christman’s model.

I believe that it is not accidental that this change in self-assessment comes only once Stevens had been challenged in his stable values, had been treated and seen differently by others, and had experienced a different enactment of roles that for him were *fixed*. Earlier in his life, Stevens had neither considered the possibility of not fully identifying with the socially available role of a butler so that it did not become suffocating nor the possibility of altering this role. The character of Miss Kenton, on the other hand, pursues both these strategies. Throughout the novel, Miss Kenton repeatedly resists full identification with her role as a housekeeper,<sup>481</sup> and she eventually decides to leave the house to get married. To be sure, as we see towards the end of the novel, the new roles she accepts are not immune to new problematic injunctions, but Miss Kenton seems to have understood more deeply the limitations of her social role as a servant. As a result of the latter, she strikes us as *less* heteronomous than Stevens.

All in all, my suggestion is that Stevens’ commitment to his role and stable values in the majority of the novel could not be seen as sufficiently critical for the purposes of assessing whether he is autonomous vis-à-vis this role in light of the injunctions of his society. What seems to be off about Stevens’ self-assessment is the impossibility, or the

---

<sup>479</sup> Ibid., 195-6.

<sup>480</sup> Ibid., 255-6.

<sup>481</sup> For example: she challenges Lord Darlington’s decisions when they conflict with her convictions (e.g. in the case of the dismissal of two Jewish maids.) (Ibid., 157)

reluctance, to place his *personal* history within a broader *social* or even *political* history. In fact, his personal history seems almost invariably the history of a *social role*, but this role is never understood within a wider context (e.g. that of post-war Britain) or questioned in light of the social configuration in which it makes sense (e.g. class structures). In this case, as I have argued, Stevens' ability to function without "short-circuits" should not be considered sufficient evidence of his autonomy when by the latter we understand an agent's capacity to assess and possibly limit their subjection.

But how are subjected individuals to achieve the kind of socio-historical self-understanding necessary to critically assess themselves and their circumstances through procedural means? Admittedly, I am proposing that individuals become, in a way, social critics, and this could be seen as setting the bar too high for autonomous agency.

In reply, let me first note that, as I have argued throughout my thesis, my proposal is not a *general* proposal of what autonomy means for *all* contexts or aims – e.g. one in which what is at stake is the legitimacy of a paternalistic intervention in a medical context or the attribution of liberal rights. Relating this comment to Stevens' example, for instance, it does in no way follow from my proposal that Stevens should be prevented from being a servant by, say, restricting his occupational choice. What I do argue is that Stevens lacks a suitable perspective on himself to determine (through a procedural test of autonomy) whether he is autonomous vis-à-vis values and character traits which have been developed as a result of oppressive circumstances. Nothing automatically follows from arguing this about the legitimacy or requirement of any paternalistic intervention in his life.

Second, the social perspective that I deem necessary to limit subjection through procedural means is demanding but it is still within the reach of subjected agents. It would be a mistake to assume that only deeply "theoretical" social critique (like the one we might tend to associate with academic contexts) could give us the tools we need to challenge settled values, problematize one's experience, and gain insight into the social roles we occupy. As

Stevens' example illustrates, I am open to the possibility of self-critique being triggered by accidental and embodied experiences. For example, Stevens' self-understanding is challenged after he is taken for someone else and sees himself enacting a different role which, as I argued, was crucial to problematize long-lasting commitments (e.g. his account of dignity). Similarly, the shock of confronting someone "like me" who enacts the same roles in a radically different way can unveil that playing (even the same) roles differently is possible.

Crucially, I want to emphasize that the social perspective that I deem necessary for a revised procedural test does not appear that demanding (intellectually or affectively) when it is thought as one which can be acquired in collaboration with others. Take feminist "consciousness raising" practices, mentioned in Chapter 3 – I argued that consciousness raising provided a good example of the kind of collective and critical hermeneutical work which might be necessary to experience some forms of alienation when oppressed and, therefore, to carry out a meaningful procedural test.

Let me analyse consciousness raising practices in a bit more detail. First, it is worth noting that these practices *are* forms of social critique.<sup>482</sup> As MacKinnon describes it:

Consciousness raising is a face-to-face social experience that strikes at the fabric of meaning of social relations between and among women and men by calling their givenness into question and reconstituting their meaning in a transformed and critical way.<sup>483</sup>

Consciousness raising thus shows that a critical social perspective on oneself and one's situation is within the reach of those oppressed. Crucially, my point is that this kind of practical social critique can contribute to the critical attitude presented in Chapter 2, insofar

---

<sup>482</sup> See for example Lorna Finlayson who refers to consciousness raising as "a kind of ideology-critique in action". The author acknowledges Katharine Jenkins for this thought. (Lorna Finlayson, *An Introduction to Feminism* (Cambridge: Cambridge University Press, 2016), 22.

<sup>483</sup> Catharine A. MacKinnon, *Toward a Feminist Theory of the State* (Cambridge: Harvard University Press, 1989), 95.

as consciousness raising involves, at the same time, the de-individualization of one's experience and the calling into question of the social identities, meanings, and conventions, which structure that experience. I unpack these claims in the paragraphs that follow:

Regarding the de-individualizing element, what I mean is that consciousness raising practices, as MacKinnon puts it, give “content and form to women's point of view”.<sup>484</sup> The self-understanding that is acquired through this kind of practice is more one on “ourselves” than on “oneself”. Practices like consciousness raising are key insofar as they provide insight on socially available spaces of possibility for character formation instead of focusing on agents' particular “personalities”.

This “women's point of view”, however, is not to be unearthed from or discovered in one's experience as it is, but needs to be critically constructed by opposing dominant social meanings and power structures. As MacKinnon notes, consciousness raising also entails interrogating social reality – e.g. “true” biological imperatives are questioned, “necessary” social conventions are shown as contingent, “men's point of view” is shown as biased and “convenient” for maintaining oppressive power relations, etc. MacKinnon claims:

The point of the process was not so much that hitherto-undisclosed facts were unearthed or that denied perceptions were corroborated or even that reality was tested, although all these happened. It was not only that silence was broken and that speech occurred. The point was, and is, that this process moved the reference point for truth and thereby the definition of reality as such. Consciousness raising alters the terms of validation by creating community through a process that redefines what counts as verification. This process gives both content and form to women's point of view.<sup>485</sup>

---

<sup>484</sup> Ibid., 87.

<sup>485</sup> Ibid., 87.

Indeed, a critical “women’s point of view” cannot be built without problematizing the biased and tainted nature of the truths which produce ‘women’. For example, MacKinnon argues that consciousness raising allows women to challenge qualities “eternally” associated with femininity (e.g. domestic skills, nurturing preferences, etc.) and to see them more overtly as “descriptions of the desired and required characteristics of particular occupants of women’s roles.”<sup>486</sup>

Crucially, then, if collective critical practices are to contribute to one’s de-subjection and are to provide opportunities for new experiences (including the experiences of alienation that are crucial for procedural tests to be meaningful), they also need to present what we learn about ourselves as contingent and open to change. In this sense, practices of social critique like consciousness raising are compatible with the Foucauldian framework presented and endorsed throughout this thesis.<sup>487</sup>

And what does “challenging ourselves” mean in practice? On the one hand, this means that, by coming to see one’s experience as shaped (e.g. enabled or obscured) by oppressive power dynamics, one may become capable of “clarifying” one’s experiences. As it emerges from the passage quoted above, MacKinnon does not present this as the main point of consciousness raising, but it is nonetheless one of the most obvious (and of the most frequently referred to) results of the process. MacKinnon acknowledges that thanks to consciousness raising “women could articulate the inarticulate, admit the inadmissible.”<sup>488</sup> Furthermore, MacKinnon notes that strong affective reactions (e.g. anger) only emerged after personal histories were seen in light of social determinations, i.e. after participants saw

---

<sup>486</sup> Ibid., 90.

<sup>487</sup> Johanna Oksala provides a Foucauldian reading of consciousness raising practices. Oksala claims: “remodelled practices of consciousness-raising would not imply simply sharing our personal stories in order to find empowering commonalities between women. The aim would rather be a problematisation of who we are and who we aspire to be—a critical reflection on the social and political conditions constitutive of our normalized experiences.” (Johanna Oksala, “In Defense of Experience,” *Hypatia* 29, no. 2 (Spring 2014): 398).

<sup>488</sup> MacKinnon, *Toward a Feminist Theory of the State*, 87.

“women’s lives as one avenue after another foreclosed by gender.”<sup>489</sup> Crucially, then, problematisation, testing, and re-interpretation of one’s own experience might enable certain affective reactions typically associated with alienation to manifest fully.

Nonetheless, as I have noted several times, the articulation of “timid” discomforts into fully-fledged experiences of alienation is only one part of what is implied when we say that a more social perspective might be necessary to experience alienation. It is important to recall that those oppressed do experience big amounts of suffering, shame, dissatisfaction, anger, frustration, and so on in their everyday lives. The problem is that these negative experiences might equally be taken as warning signs of, say, one’s dissatisfaction given one’s social position, or as evidence that one needs to keep working on oneself to meet social standards which one may have never questioned. In other words, subjected agents may “feel” all the negative feelings and emotions typically associated with alienation but still fail to feel alienated from a particular social role or social situation. So, how come some “negative” experiences succeed in disrupting subjected self-management while others merely reinforce it or “crush” subjects altogether? I have suggested that a key element to explain this difference is whether agents inform their self-assessments by way of a social perspective. Feminist consciousness raising practices are also a good example in this regard:

For instance, those living in oppressive social conditions may be overwhelmed or “disempowered” by negative experiences or affects caused by, e.g., oppressive social stereotypes. Feelings of shame or anger may affect the way in which one relates to oneself by, for example, undermining one’s self-confidence. As I noted in Chapter 4, when analysing Benson’s proposal, this is a risk that needs to be addressed by a theory of autonomy. My suggestion was that relating to one’s suffering with an understanding of the fact that this

---

<sup>489</sup> Ibid., 91.

suffering does not derive from personal faults (or, even, acknowledging that I am not the only one suffering for the same reasons) could help to mitigate “disempowering” effects.

Indeed, this potential is also confirmed by MacKinnon’s analysis. MacKinnon notes that consciousness raising has “unburdening” effects: seeing one’s suffering as originating from the dictates of “powerful social conventions often makes women feel unburdened, since individual failures no longer appear so individualized.”<sup>490</sup>

Furthermore, the de-individualisation of suffering may also play, I believe, an important role in avoiding a potentially “paralysing” effect of my proposal. Admittedly, asking agents to challenge or “refuse who they are” may come at very high psychological costs. As I discussed in Chapter 3, agents attach to their social roles because these roles, to put it roughly, allow them to exist in a social space. This attachment could not merely be undone by “understanding” that, say, a social identity has been contingently constructed within an oppressive structure.<sup>491</sup> In reply, let me provide two brief comments:

First, as showed by my efforts to secure the conditions of possibility for experiencing alienation (and, furthermore, by my efforts to distinguish experiences of alienation from “crushed” agency), I am fully aware of the limitations of a model that reduces de-subjection to an intellectual task or to a cognitive “understanding”. Experiences of alienation, as I have argued in the preceding chapters, do not merely work on a “cognitive” level.

Second, I want to reiterate that my aim has not been to present the task of limiting heteronomy as one in which oppressed subjects become less oppressed by renouncing their possibilities to exist in a social space. My suggestion, instead, is that the processes of critical self-assessment which I have described throughout my thesis provide more than the possibility to see one’s way of being as contingent, or to envisage other ways of being.

---

<sup>490</sup> Ibid.

<sup>491</sup> See for example Allen, *The Politics of Our Selves*, 80 for a similar argument on the insufficiency (for practical or political purposes) of simply challenging what we *think* about gender.

Individuals also place themselves within a “we”, which may need to be challenged in important ways to limit one’s subjection, but which could also provide a base for a critical “community” (as MacKinnon notes).<sup>492</sup> That is, individuals also become aware that being otherwise, with others, is possible.

### III. Conclusion

I have already summarized the key points of my argument in the introduction to this chapter and I have considered possible objections to my proposal (and potential replies to those objections) in the previous section, so I keep my conclusion short to avoid unnecessary repetition.

My aim in this thesis has been to critically assess existing models of personal autonomy in light of their capacity to deal with oppression. I have done so by looking into paradigmatic examples of the two main strategies available in the literature on autonomy: Dworkin’s, Mele’s, and Christman’s *procedural* accounts; and Stoljar’s, Benson’s, and Oshana’s *substantive* accounts.

I have mostly focused on assessing existing models in light of the challenges posed by ‘subjection’, namely a (common) form of oppression which cannot be understood simply as a way of coercing, repressing, deceiving, or manipulating individuals of a certain “kind”. As I have argued throughout this thesis, ‘subjection’ *produces* subjects who self-relate and conduce themselves qua subjects of certain kinds. Crucially, then, some forms of self-management might sustain or reinforce one’s oppression. I have argued that some cases of

---

<sup>492</sup> To be sure, as I have acknowledged throughout this thesis, no social space will be entirely free from power dynamics and this is why a critical attitude needs to be thought as an open-ended task and autonomy cannot be thought as an all-or-nothing matter. Still, as I have argued, some alternatives and strategies do significantly *limit* one’s oppression.



gender, race, and class oppression – where one’s development as a whole is shaped in light of a sexist, racist, or exploitative structure – constitute examples of oppression as subjection.

I have argued that existing substantive and procedural models are, as they stand, ill-suited to account for and to tackle subjection. On the one hand, procedural accounts seem to assume a limited understanding of oppression, namely one in which oppression is mostly a phenomenon which *disrupts* individuals’ development. As a result, I have argued, procedural models lack resources to problematize settled characters and values.

Substantive accounts, on the other hand, tend to be much more sensitive to structural oppression, and to the influences which shape one’s development, value-framework, and emotional tendencies. However, available substantive accounts tend to deduce, from the fact that agents’ psychologies are shaped or constrained by oppression, the conclusion that they are irredeemably so, unless external change happens.

To challenge the latter conclusion, I have argued that ‘subjection’ cannot be reduced to the full and one-off ‘internalisation’ of oppression (e.g. to the internalisation of an oppressive moral code or stereotype). ‘Subjection’ involves a complex (and continuous) combination of “external” and “internal” influences. For example, it involves both the training of bodies and the structuring of public space *and* injunctions to self-relate in specific ways in light of socially-defined “truths” about who one is. This complex view of oppression may seem daunting at first glance but, as I have shown, this complexity also makes the opportunities for pushing back on one’s subjection much more abundant than the model of ‘internalisation’ may lead us to think. Available substantive models, nonetheless, cannot sufficiently account for the subjective strategies which may make agents who live under the *same* (normative or social) conditions, less subjected.

However, I have acknowledged that, in some particularly challenging cases, subjective strategies to limit oppression will not take us far enough. The latter is the case, I have claimed, when oppression is maintained mainly through “repressive” means. That is, in the latter

cases, what sustains oppression is not typically the way in which free individuals conduct themselves, relate to themselves, or “who” and “how” they consider themselves to be. Rather, in these more challenging cases, selves might be “crushed”, spaces of possible action are severely restricted, bodies are subjected to coercion or sheer violence, or power relations are so asymmetrical that subjects are, in a way, “powerless” to significantly alter their situation. In these cases, significant external and/or structural change must be brought about before an agent could have the necessary control over her life that is needed for self-government. My proposal, then, incorporates a substantive condition: ‘non-domination’.

Therefore, instead of overcoming the difficulties faced by existing (procedural and substantive) models of autonomy through a single strategy, I have argued that, since oppression can take different forms, it is necessary to adopt different strategies depending on the kind of oppression at stake. When agents are dominated, heteronomy should be decided substantively. When agents are subjected, however, limiting heteronomy through a *revised* procedural strategy is possible.

The revisions necessary to strengthen the procedural account which I have deemed the most promising (i.e. Christman’s historical account) are mainly of two kinds:

First, in contexts of subjection, becoming more autonomous requires us to go *against* ‘authenticity’ (as defined in procedural models). I have argued that any attempt to limit subjection will be seriously compromised by injunctions to “remain true” to who one “really” is, especially when the latter is defined by reference to one’s settled character or sustained commitments. As I have shown in Chapter 2, challenging the notion of ‘authenticity’ also entails challenging very narrow ideas of ‘reflective competence’ which promote coherence (internal or diachronic) over self-questioning, or “self-discovery” over self-transformation.

Second, I have defended the position that to limit heteronomy in contexts of subjection, agents need to be able to place their personal histories within a wider history which is social and that concerns the spaces of possibility for character formation. A minimal

awareness of the social utility of certain social roles is crucial if agents are to have meaningful opportunities to challenge their cognitive and affective endorsement of those roles. For example, not experiencing alienation from a social role, when this happens without a minimal understanding of the fact that this role makes sense (and contributes to) a certain social hierarchy, is insufficient.

All in all, a wider social perspective on oneself, made possible by acknowledging that personality and character formation happen within a social space of possibility structured by (frequently) oppressive circumstances, may allow for a critical procedural self-assessment in contexts of subjection. Crucially, this self-understanding pushes back on a narrative that presupposes that settled ways of being, recurrent affective responses, and habitual desires, are necessarily “one’s own”.

## Bibliography

- Alfred Mele. *Autonomous Agents*. Oxford: Oxford University Press, 2001.
- Allen, Amy. "Feminism, Foucault, and the Critique of Reason: Re-reading the History of Madness." *Foucault Studies*, no. 16 (September 2013): 15-31.
- . "Power and the Politics of Difference: Oppression, Empowerment, and Transnational Justice." *Hypatia* 23, no. 3 (July-September 2008): 156-172.
- . "Recognizing Domination: Recognition and Power in Honneth's Critical Theory." *Journal of Power* 3, no.1 (April 2010): 21-32.
- . *The politics of Our Selves*. New York: Columbia University Press, 2008.
- Appiah, Kwame A. *The Ethics of Identity*. New Jersey: Princeton University Press, 2005.
- Bartky, Sandra L. *Femininity and Domination*. New York: Routledge, 1990.
- Barvosa-Carter, Edwina. "Mestiza Autonomy as Relational Autonomy: Ambivalence & the Social Character of Free Will." *The Journal of Political Philosophy* 15, no. 1 (2007): 1-21.
- Bennice, Jennifer and Resick, Patricia. "Marital Rape: History, Research, and Practice." *Trauma, Violence, & Abuse* 4, no. 3 (July 2003): 228-246.
- Benson, Paul. "Autonomy and Oppressive Socialization." *Social Theory and Practice* 17, no. 3 (Fall 1991): 385-408.
- . "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 S. Taylor, 124-142. Cambridge: Cambridge University Press, 2005.
- . "Free Agency and Self-Worth." *The Journal of Philosophy* 91, no. 12 (December 1994): 650-668.
- . "Freedom and Value." *The Journal of Philosophy* 84, no. 9 (September 1987): 465-486.
- Bevir, Mark. "What is Genealogy?" *Journal of the Philosophy of History* 2 (2008): 263-275.
- Butler, Judith. *The Psychic Life of Power: Theories in Subjection*. Stanford: Stanford University Press, 1997.

- Chega de Fiu Fiu (Enough with Catcalling)*, directed by Amanda Kamanchek and Fernanda Frazão (2018; Brazil: Brodagem Filmes).
- Christman, John. "Autonomy and Liberalism: A Troubled Marriage?" In *The Cambridge Companion to Liberalism*, edited by Steven Wall, 141-162. Cambridge: Cambridge University Press, 2015.
- . "Autonomy and Personal History." *Canadian Journal of Philosophy* 21, no. 1 (March 1991): 1-24.
- . "Coping or Oppression." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 201-226. Oxford: Oxford Scholarship Online, 2014.
- . "Decentered Social Selves: Interrogating Alienation in Conversation with Rahel Jaeggi". In *From Alienation to Forms of Life: The Critical Theory of Rahel Jaeggi*, edited by Amy Allen and Eduardo Mendieta, 41-58. Pennsylvania: The Pennsylvania State University Press, 2018.
- . "Telling Our Own Stories: Narrative Selves and Oppressive Circumstance." In *The Philosophy of Autobiography*, edited by Christopher Cowley, 122-140. Chicago: The University of Chicago Press, 2015.
- . *The Politics of Persons*. Cambridge: Cambridge University Press, 2009.
- Cremonesi, Laura et al., eds. *Foucault and the Making of Subjects*. Lanham: Rowman and Littlefield International, 2016.
- Cudd, Ann E. "Adaptations to Oppression: Preference, Autonomy, and Resistance." In *Personal Autonomy and Social Oppression*, edited by Marina Oshana, 142-160. New York: Routledge, 2015.
- . *Analysing Oppression*. Oxford: Oxford University Press, 2006.
- Davidson, Arnold I. "Ethics as Ascetics: Foucault, the History of Ethics, and Ancient Thought." In *Cambridge Companion to Foucault*, edited by Gary Gutting, 123-148. Cambridge: Cambridge University Press, 2005.
- Dreyfus, Hubert L. and Rabinow, Paul. *Michel Foucault: Beyond Structuralism and Hermeneutics*. Chicago: The University of Chicago Press, 1983.

- Dworkin, Gerald. "Autonomy and Behavior Control" *The Hastings Center Report* 6, no. 1 (February 1976): 23-28.
- . "The Concept of Autonomy." In *The Inner Citadel: Essays on Individual Autonomy*, edited by John Christman, 54-62. Oxford: Oxford University Press, 1989.
- . *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press, 1988.
- Ehrenberg, Alain. *The Weariness of the Self: Diagnosing the History of Depression in the contemporary Age*. Montreal: McGill-Queen's University Press, 2010.
- Fernández Huidobro, Eleuterio and Rosencof Mauricio. *Memorias del Calabozo*. Navarra: Ed. Txalaparta, 1993.
- Finlayson, Lorna. *An Introduction to Feminism*. Cambridge: Cambridge University Press, 2016.
- . "How to Screw Things with Words." *Hypatia* 29, no. 4 (Fall 2014), 774-789.
- Foucault, Michel. *About the beginning of the hermeneutics of the self: Lectures at Dartmouth College, 1980*. Chicago: The University of Chicago Press, 2015.
- . *An Introduction, Vol. 1 of the History of Sexuality*. Translated by Robert Hurley. New York: Pantheon Books, 1978.
- . *Discipline and Punish: The Birth of the Prison*. Translated by Alan Sheridan. New York: Vintage Books, 1995.
- . "Friendship as a Way of Life." In *Ethics, Subjectivity, and Truth, Vol 1 of The Essential Works of Michel Foucault 1954-1984*, edited by Paul Rabinow, 135-140. New York: The New York Press, 1997.
- . "Governmentality." In *Power Vol. 3 of Essential Works of Michel Foucault 1954-1984 Vol. 3*, edited by James D. Faubion, 201-222. New York: The New York Press, 2000.
- . *Histoire de la folie à l'âge Classique*. Paris: Gallimard, 1972.
- . *History of Madness*. Translated by Jonathan Murphy and Jean Khalfa. New York: Routledge, 2006.

- . “Nietzsche, la Généalogie, l’Histoire.” In Bachelard Suzanne et al. (eds.) *Hommage à Jean Hyppolite*, 145-172. Paris: Presses Universitaires de France, 1971.
- . “Omnes et Singulatim: Toward a Critique of Political Reason.” In *Power Vol. 3 of Essential Works of Michel Foucault 1954-1984 Vol. 3*, edited by James D. Faubion, 298-325. New York: The New York Press, 2000.
- . “On the Genealogy of Ethics: An Overview of Work in Progress.” In *Ethics, Subjectivity, and Truth, Vol 1 of The Essential Works of Michel Foucault 1954-1984*, edited by Paul Rabinow, 253-280. New York: The New York Press, 1997.
- . “Sex, Power, and The Politics of Identity.” In *Ethics, Subjectivity, and Truth, Vol 1 of The Essential Works of Michel Foucault 1954-1984*, edited by Paul Rabinow, 163-174. New York: The New York Press, 1997.
- . *Society Must be Defended: Lecture at the Collège de France 1975-6*. Translated by David Macey. New York: Picador, 2003.
- . “The Ethics of the Concern of the Self as a Practice of Freedom.” In *Ethics, Subjectivity, and Truth, Vol 1 of The Essential Works of Michel Foucault 1954-1984*, edited by Paul Rabinow, 281-301. New York: The New York Press, 1997.
- . “The minimalist Self.” In *Foucault, Michel, Interviews and other writings 1977-1984*, edited by Lawrence D. Kritzman, 3-16. London: Routledge, 1988
- . “There Can’t be Societies without Uprisings.” In *Foucault and the Making of Subjects*, edited by Laura Cremonesi et al., 25-51. Lanham: Rowman and Littlefield International, 2016.
- . “The Subject and Power.” In *Power Vol. 3 of Essential Works of Michel Foucault 1954-1984*, edited by James D. Faubion, 326-348. New York: The New York Press, 2000.
- . “Truth and Power.” In *Power Vol. 3 of Essential Works of Michel Foucault 1954-1984*, edited by James D. Faubion, 111- 133. New York: The New York Press, 2000.
- . “What is Enlightenment?” In *Ethics, Subjectivity, and Truth, Vol 1 of The Essential Works of Michel Foucault 1954-1984*, edited by Paul Rabinow, 303-319. New York: The New York Press, 1997.

- . *Psychiatric Power: Lectures at the College de France 1973-74*. Translated by Graham Burchell. Basingstoke: Palgrave Macmillan, 2006.
- . “*Qu’est-ce que la Critique*” suivi de “*La Culture de Soi*,” edited by Henri-Paul Fruchaud and Daniele Lorenzini. Vrin: Paris, 2015
- . *The Government of Self and Others, lectures at the Collège de France 1982-1983*. Hampshire: Palgrave Macmillan, 2010.
- . *The Use of Pleasure: Vol. 2 of The History of Sexuality*. New York: Vintage Books, 1990.
- Frankfurt, Harry J., “Freedom of the Will and the Concept of a Person.” In *The Inner Citadel: Essays on Individual Autonomy*, edited by John Christman, 63-76. New York: Oxford University Press, 1989.
- . *The Importance of What We Care About*. Cambridge: Cambridge University Press, 1988
- Fraser, Nancy. *Justice Interruptus: Critical Reflections on the Postsocialist Condition*. New York: Routledge, 1997.
- Freyenhagen, Fabian and O’Shea Tom. “Hidden substance: mental disorder as a challenge to normatively neutral accounts of autonomy.” *International Journal of Law in Context* 9, Special Issue 01 (March 2013): 53-70.
- Freyenhagen, Fabian. “Autonomy’s Substance.” *Journal of Applied Philosophy* 34, no. 1 (February 2017): 114-129.
- Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford Scholarship Online, 2007.
- Friedrich Nietzsche. *On the Genealogy of Morals*, trans. W. Kaufmann. New York: Vintage Books Editions, 1989.
- Frye, Marilyn. *Oppression*. In *Gender Basics: Feminist Perspectives on Women and Men*, edited by Anne Minas, 10-16. Belmont: Wadsworth Publishing, 2000.
- Galli, Gabriel. “Em país hostil a travestis, brasileira faz arte a partir de sua identidade e denuncia transfobia.” *Global Voices*, July 6, 2020.



<https://pt.globalvoices.org/2020/07/06/em-pais-hostil-a-travestis-brasileira-faz-arte-a-partir-de-sua-identidade-e-denuncia-transfobia/> (last accessed July 2020)

- Geuss, Raymond, "Genealogy as Critique." *European Journal of Philosophy* 10, no. 2 (December 2002): 209-215.
- Guillaume le Blanc. "Becoming a Subject in Relation to Norms." In *Foucault and the Making of Subjects*, edited by Laura Cremonesi, 129-136. Lanham: Rowman and Littlefield International, 2016.
- Hacking, Ian. *Historical Ontology*. Cambridge: Harvard University Press, 2002.
- Han, Beatrice. *Foucault's Critical Project: Between the Transcendental and the Historical*. Stanford: Stanford University Press, 2002.
- Han-Pile, Béatrice. "Foucault, normativity and critique as a practice of the self." *Continental Philosophy Review* 49, no. 1 (2016): 85-101.
- Hartman, Saidiya V. *Scenes of Subjection: Terror, Slavery, and Self-making in Nineteenth-Century America*. New York: Oxford University Press, 1997.
- Holt, Thomas. "Pouvoir, savoir et race : À propos du cours de Michel Foucault 'Il faut défendre la société'." In *Lectures de Michel Foucault, Vol. 1*, edited by Jean-Claude Zancarini, 81-96. Paris: ENS Editions, 2001.
- Illouz, Eva. *Consuming the Romantic Utopia: Love and the Cultural Contradictions of Capitalism*. Berkeley: University of California Press, 1997.
- Ishiguro, Kazuo. *The Remains of the Day*. London: Faber and Faber, 1989.
- Jütten, Timo. "Sexual Objectification." *Ethics*, 127 (October 2016): 27-49.
- Lloyd, Genevieve, "The Man of Reason." *Metaphilosophy* 10, no. 1 (January 1979): 18-37.
- Lugones, María. *Pilgrimages/Peregrinajes: Theorizing Coalition Against Multiple Oppressions*. New York: Rowman & Littlefield Publishers, 2003
- Luker, Kristin. "Contraceptive Risk Taking and Abortion: Results and Implications of a San Francisco Bay Area Study." *Studies in Family Planning* 8, no. 8 (Aug., 1977): 190-196.

- . *Taking Chances: Abortion and the Decision Not to Contracept*. Berkeley: University of California Press, 1975.
- MacKinnon, Catharine A. *Toward a Feminist Theory of the State*. Cambridge: Harvard University Press, 1989.
- MacPhee, Graham. *Postwar British Literature and Postcolonial Studies*. Edinburgh: Edinburgh University Press, 2011.
- McNay, Lois. "Self as Enterprise: Dilemmas of Control and Resistance in Foucault's 'The Birth of Biopolitics'." *Theory, Culture, and Society* 26, no. 6 (2009): 55–77.
- McWhorter, Ladelle. "Post-Liberation Feminism and Practices of Freedom." *Foucault Studies*, no. 16 (September 2013): 54-73.
- Meyers, Diana T. *Being yourself: Essays on Identity, Action, and Social Life*. New York: Rowman & Littlefield Publishers, 2004.
- . "Feminist Debates Over Values in Autonomy Theory." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 114-140. Oxford: Oxford University Press, 2014.
- Noggle, Robert. "Autonomy and the Paradox of Self-Creation: Infinite Regress, Finite Selves, and the Question of Authenticity." In *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy*, edited by James S. Taylor, 87-108. Cambridge: Cambridge University Press, 2005.
- O'Shea, Tom. "Autonomy and Orthonomy." *Journal of Moral Philosophy*, 12 (2015): 619-637.
- Oksala, Johanna. "In Defense of Experience." *Hypatia* 29, no. 2 (Spring 2014): 388-403.
- Oshana, Marina. "A Commitment to Autonomy is a Commitment to Feminism." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 141-160. Oxford: Oxford University Press, 2014.
- . "Autonomy and the Question of Authenticity." *Social Theory and Practice* 33, no. 3 (July 2007): 411-29.

- . “Is Social-Relational Autonomy a Plausible Ideal?” In *Personal Autonomy and Social Oppression: Philosophical Perspectives*, edited by Marina Oshana, 3-24. New York: Routledge, 2015.
- . *Personal Autonomy in Society*. Aldershot: Ashgate, 2006.
- , ed. *Personal Autonomy and Social Oppression*. New York: Routledge, 2015.
- Pereira, Gustavo. *Imposed Rationality and Besieged Imagination Practical Life and Social Pathologies*. Cham: Springer, 2019.
- Saar, Martin. “Understanding Genealogy: History, Power, and the Self.” *Journal of the Philosophy of History* 2 (2008): 295–314.
- Salberg, Jill. “The Texture of Traumatic Attachment: Presence and Ghostly Absence in Transgenerational Transmission.” *The Psychoanalytic Quarterly* 84, no. 1 (2015): 21-46.
- Schaub, Jörg. “Aesthetic Freedom and Democratic Ethical Life: A Hegelian Account of the Relationship Between Aesthetics and Democratic Politics,” *European Journal of Philosophy* (2018): 1-23.
- Srinivasan, Amia. “Does Anyone Have the Right to Sex?” *London Review of Books*, March 22, 2018.
- . “Genealogy, Epistemology and Worldmaking.” *Proceedings of the Aristotelian Society* 119, Part 2 (2019): 127-156.
- Stoljar, Natalie, “Autonomy and Adaptive Preference Formation.” In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 228-252. Oxford: Oxford University Press, 2014.
- . “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.
- . “Living Constantly at Tiptoe Stance: Social Scripts, Psychological Freedom, and Autonomy.” In *Personal Autonomy and Social Oppression: Philosophical Perspectives*, edited by Marina Oshana, 105-123. New York: Routledge, 2015.

- Taylor, James S. (Ed.) *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy*. Cambridge: Cambridge University Press, 2005.
- Westlund, Andrea C. "Rethinking Relational Autonomy." *Hypatia* 24, no. 4 (2009): 26-49.
- Worsdale, Rosie, "Recognition, Ideology, and the Case of "Invisible Suffering." *European Journal of Philosophy* 26 (2018): 614–629.
- Young, Iris M. *Justice and the Politics of Difference*. Princeton: Princeton University Press, 1990.
- Zutlevics, Tamara L. "Towards a Theory of Oppression." *Ratio* 15 (March 2002): 80-102.