DEFENDING THE COHERENCE AND PRACTICABILITY OF AUTONOMY THROUGH A
MULTI-LEVEL ANALYTICAL APPROACH

JAMIE ROBERTSON

A DISSERTATION SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PHILOSOPHY

YORK UNIVERSITY

TORONTO, ONTARIO

November 2018

Abstract

The objective of this dissertation is to develop a coherent account of autonomy that builds on the general understanding of autonomy as involving deciding or discovering for oneself what is valuable and living accordingly. I will advance a multi-level, multi-factor theory of autonomy while responding to potential criticisms relating to autonomy's coherence as a concept and practicability as a capacity. In addition to solving problems relating to coherence and practicability, my project makes several contributions. These include providing a rich account of autonomy thresholds, revealing a middle ground between causally and constitutively relational autonomy, and offering a novel conception of autonomy that has implications for our social and political responses in instances where people are deemed not to meet particular autonomy thresholds.

In my first chapter I refute allegations that taking practical considerations into account in developing a theory of autonomy constitutes a wrongful inclusion of normative considerations into what should be a purely conceptual analysis. I also respond to situationist arguments against the possibility of autonomy. In so doing, I will articulate the common-sense psychological standard I will use to judge theoretical adequacy throughout the remainder of the dissertation.

In the second chapter, I track how common-sense concerns about the practicability of autonomy have been used to bring contemporary conceptions of autonomy more in line with human experience and limitations. I argue that while considerable nuance has been added to the otherwise proceduralist picture of autonomy, this increased complexity exacerbates concerns about the (lack of) conceptual coherence of autonomy and raises concerns that the exercise of autonomy is overly demanding.

In the third chapter I respond to Nomy Arpaly's claim that the concept of autonomy is incoherent. I do so by advancing a three-level approach to analyzing autonomy in which important elements of Arpaly's discussion of moral responsibility feature at different levels of analysis. My approach provides a different perspective on Arpaly's critique of autonomy and opens new avenues of inquiry into the relationship between autonomy and moral responsibility. While my model helps join different aspects of autonomy together into a coherent picture, it simultaneously reveals the extent to which the exercise of autonomy requires an *extensive* range of abilities and is highly *complex*. At the end of this chapter a common-sense concern emerges: we may not expect autonomy to be widely practicable.

Defending autonomy against these renewed concerns about practicability will be the objective of the final chapters of the dissertation. This defense will rely on three additional features of my theory of autonomy: degree, automaticity, and reinforcement/substitution. In both chapters four and five, I will endeavour to convince the reader that these proposed features are plausible on a common-sense understanding of human psychology.

Acknowledgements

My warmest thanks to Alice MacLachlan and Bob Myers for their supervision, and to Mike Giudice and Louis-Philippe Hodgson for their guidance. I am grateful for your attention to, and collegial scrutiny of, my work in support of the completion of this project.

Table of Contents

Introduction

This project offers, at least in spirit, a response to John Christman's call, issued in his introduction to *The Interior Citadel* (1989), for philosophers to take on the task of identifying the conceptual core of autonomy. Autonomy is often, if not generally, thought to be associated with self-governance (Levy 2006, 7, Meyers 1989, 13-4, Killmister 2013, 95, Christman 1991, 1), but what it means to be self-governing, and how one comes to qualify as such, remains in need of significant elucidation (Killmister 2013, 95). As a preliminary step toward offering a more substantive broad idea of autonomy from which to work, I will take autonomy to involve something like "deciding [or discovering] for oneself what is valuable, and living one's life in accordance with that decision" (Colburn 2011, 8). On this understanding, autonomy plays an important role in the adoption and pursuit of personal values (Christman 1989, 19) since it is through the process of exercising autonomy that various aspects of our lives come to count as our own in an important and valuable sense (Levy 2006, 2, Mackenzie 2014, 18). At the time of Christman's call, and in the decades since, a significant proportion of philosophical work relating to autonomy has aimed at the development of a theoretically adequate and unified articulation of autonomy as a concept. Some think that these efforts "have not been particularly successful" (Anderson 2014, 356), as evidenced by the persistent lack of agreement about the basic aspects of autonomy (Mackenzie 2014, 18) and the complaints in some circles that the meaning of 'autonomy' is "at least as elusive and complex as the natural-language terms it was supposed to help clarify" (Arpaly 2002, 118-19).

I do not think, however, that we need to be quite so cynical about the potential for autonomy to be an intelligible and useful philosophical concept. But, this is not to say that articulating a coherent conception of autonomy is without difficulty. There are several challenges that must be

faced when undertaking such a task. As I see it, theorizing about autonomy faces at least four challenges.

First, historically there has not been consistent terminology for discussing the process by which people direct their lives on the basis of the things they deem to be of value. At times autonomy has become wrapped up with freedom and independence of various kinds (Christman 1989, 4), with liberty, or with responsibility. The effect of this terminological inconsistency has been to obscure the sense in which autonomy is a distinct concept. A coherent conception of autonomy should resist this tendency for conflation with other concepts, but should also embrace relevant insights presented in non-autonomy terms.

Second, a coherent conception must navigate the considerable variety of additional concepts that are invoked in discussions about autonomy. These range widely in nature. They stretch between concepts related to political rights and status, to those related to virtues and skills, and again to concepts relating to attitudes and mental states. A coherent conception of autonomy should provide some account of how so many of these ideas relate to each other within the concept of autonomy.

The third problem is that different authors take different stances on the object(s) to which the labels 'autonomy' or 'autonomous' can rightly be applied and in what way. By this I mean that 'autonomy' is variously applied to people, to choices, to actions, to goals, values, or beliefs, and at times to entire lives. Furthermore, 'autonomy' might be applied to them in the sense of an ideal, a capacity, a condition, or a right (Feinberg 1989, 28, Gracia 2012, 57). And, the application may be either global or local in scope, as well as construed as a binary property or as a matter of degree. A coherent conception of autonomy must provide guidance on how to approach these diverse usages.

Fourth, and finally, there is the challenge related to ensuring autonomy as a concept is sensitive to the realities of human experience. If a concept of autonomy is so demanding as to exclude all, or even most, people then I will argue we can rightfully doubt its acceptability as an adequate account of the concept.

The objective of this dissertation is to develop a coherent account of autonomy centred around the general idea of deciding for oneself what is valuable and living accordingly, while grappling as much as possible with these four challenges. I offer here some preliminary remarks regarding my approach to each of these.

From the general definition of autonomy as involving deciding or discovering what is of value and living accordingly, it seems to me to follow that autonomy is a capability that is exercised by people. However, the general conception from which I am working also suggests that the exercise of autonomy consists of the exercise of certain subordinate functions, namely those functions involved in the activity of deciding for oneself what is important, and those involved in living according to what one has deemed important. While these might superficially map onto the distinction well recognized in contemporary philosophical literature between autonomy's authenticity and competency conditions (Mackenzie 2014, 18), the conception I will develop over the course of the following pages will complicate this distinction somewhat.[1] In addition, I will agree with Catriona Mackenzie that trying to account for autonomy through the specification of necessary and sufficient conditions is not likely to be successful (Mackenzie 2014, 18).[2] But, whereas Mackenzie proposes that the concept of autonomy be divided up into three dimensions

---

[1] Especially insofar as this distinction might be understood as distinguishing between internal and external aspects of autonomy. Part of my project will involve blurring this distinction.
[2] Which is something Gerald Dworkin proposed decades ago (Dworkin 1981).

or axes reflecting different sets of considerations relating to autonomy, I defend a three-level analysis of autonomy in which there are dynamic relationships between the levels.

The details of the view will be developed in greater depth as the dissertation proceeds, but the basic picture is something like the following.

At the highest level of analysis, the *meta* level, autonomy is characterized by what we take the outcomes or implications of the exercise of autonomy to be. Thus, the aspects of our concept of autonomy at this level are not themselves *part* of the exercise of autonomy, but rather about those who can exercise autonomy. At this level we might discuss what virtues or characteristics we think autonomous people will have, or assign people who exercise autonomy satisfactorily certain forms of social recognition or political status. At the intermediate level of analysis, the *meso* level, I take autonomy to be characterized by the functions that would be required to make authentic decisions about what is valuable and live according to those decisions. Thus, these functions are firmly part of the exercise of autonomy. A person will warrant the lobal 'autonomous' based on an assessment of their ability to satisfactorily perform autonomy functions.[3] Meta-level descriptions will apply to people who are deemed to be autonomous. Thus, there is a clear connection between the meso- and meta- levels of analysis which is characterized by facts at the meso-level providing the basis for meta-level claims. Finally, at the lowest level of analysis, the *micro* level, autonomy is characterized by the means by which the functions of the *meso* level are (or can be) performed. The means include things like skills, capacities and self-regarding attitudes—but also other elements that are less obvious.

---

[3] While I think that local assessments of autonomy are basic, I expect that only some meta-level descriptions can be appropriately applied based on local assessments of functioning, whereas others can only be applied based on a more global ability to function adequately.

One of the hallmark features of my view of autonomy is that in different people, and in different contexts, the *meso* level functions can be achieved by different means. The reader will find that this has implications for whether we should reasonably expect to be able to specify necessary and sufficient conditions for autonomy. In short, I do not believe we can have such an expectation, but my defense of this point will have to wait until the final chapter of this work. However, if a unitary concept is one for which it is possible to specify necessary and sufficient conditions (Mackenzie 2014, 17), then clearly I do not think that autonomy is a unitary concept. Nonetheless, the schema above should give the reader a sense of how I think that the different ideas involved in the concept of autonomy can be *unified*, that is, through their participation in this conceptual *structure* in which the different elements help explain and account for each other.

This basic picture hints not only at my views on how autonomy is a coherent concept (challenge 2), but also suggests the nature of my stance on the third challenge faced in characterizing autonomy (the one relating to which objects can rightly be described by the terms autonomy and autonomous, and in what way). However, since I will not much discuss my stance on the questions involved in the third challenge elsewhere, I will pause here to outline what hints the reader should take from the sketch of my view of autonomy provided above.

Given the general understanding of the phenomenon of autonomy I articulated above, my central interest is in autonomy as a capacity exercised by people. We call people autonomous when they are able to perform the functions required to decide for themselves what is valuable and to live accordingly. The application of the terms 'autonomy' or 'autonomous' to actions and beliefs is, in my view, best understood as a form of shorthand by which we really mean to say that a person's action follows from (or is an instance of) the satisfactory exercise of autonomy functions. Our concern in many cases is whether beliefs, values, and the like are *authentic*, by

which we mean whether a person arrived at them or endorsed them as the result of a satisfactory

exercise of the subset of functions related to deciding for oneself. I would say that claims about

whether a person is autonomous, or their life is autonomous, would be based on making a global

judgement about their ability to satisfactorily exercise the *meso* level functions. Assessments of

meso-level functioning in particular situations will play an important role in the theory I am

articulating, but we might also make more generalized judgements about a person across

situations or across time. Thus, my answer to the question about what *objects* autonomy should

be applied to is people, and they should be assigned that label if they perform the *meso* level

functions adequately primarily in particular (local) circumstances.

We can make sense of the other usages from the starting point of the person and their level of

performance, though in these cases 'autonomy' is assigned as a condition rather than referring to

a capacity. I will attempt to be consistent in qualifying that the exercise of the relevant functions

should be *adequate* or *sufficient*, and this should suggest to the reader that on my view autonomy

is a matter of degree. Thus when we say someone has acted or decided autonomously, we are not

saying that they have exercised the relevant functions perfectly, but that they have passed some

threshold or standard. But, I will argue these thresholds will vary in demandingness in particular

instances and contexts, or relative to specific questions. A fuller articulation of my position will

be provided in Chapter 1, but in essence I take autonomy thresholds to be shaped by human

limitations as well as the variation of stakes, pretensions, and aspirations depending on the

question and context. We should expect that what counts as an adequate exercise of the *meso*

level functions will shift relative to these factors. Thus, my stance on the remaining aspects of

the third challenge in characterizing autonomy is that judgements about autonomy are in the first

place local rather than global, and are a matter of degree with thresholds that vary based on contextual factors.

In the earlier sections of the dissertation, I will articulate and outline responses to some common conceptual problems that autonomy theorizing has encountered as part of contending with terminological instability (the first challenge) in the process of developing a coherent conception of autonomy. In this phase I will not always critique the (often historical) views I present on all contentious points. This is because I intend them as examples of particular theoretical issues that have needed to be resolved. Thus, while I will not give a comprehensive treatment of the historical uses of the term 'autonomy', rehearse the list of terms to which it has been similar in meaning, or give comprehensive counter arguments to the positions I discuss, the reader should not take my silence on certain problems with these uses as acquiescence to, or acceptance of, them. My treatment of the first challenge is somewhat indirect, and my treatment of the third challenge is brief, because I take the challenges of unifying disparate ideas (challenge 2) and establishing practicability (challenge 4) as the most difficult and important parts of the project. The former challenge involves developing an account of autonomy in which the *extensiveness* and *complexity* of the concept can be managed in a way that preserves or enhances our sense that this concept is coherent and theoretically fruitful. The distinction between *extensiveness* and *complexity* will be useful going forward. Briefly, I take *extensiveness* to refer to the number of ideas that participate in the conceptualization of autonomy, whereas *complexity* refers to the relationships between these various ideas.[4] The latter challenge relates to ensuring that the

---

[4] Thanks to A. MacLachlan for this point.

implications of the concept do not either make it impossible to live up to or cause it to be arbitrarily exclusionary.

My project, and the process of addressing these two main challenges, in the service of developing a way of interpreting autonomy that is both coherent and practicable, will proceed in five stages.

In Chapter 1, I will take on the question of whether we should really be concerned with practicability when attempting to refine our understanding of autonomy. I will reject views that claim that taking practical considerations into account in developing a theory of autonomy constitutes a wrongful inclusion of normative considerations into what should be a purely conceptual analysis. In the process of making this argument I will advance my positive view of how contextual features shape autonomy thresholds. I will also argue that practicability is a legitimate consideration in theorizing about autonomy because my account of the exercise of autonomy appeals to cognitive and other functions. Having established this, I defend the practicability of autonomy against some preliminary objections. Against the background of my refutation of situationist arguments against autonomy, I will articulate the common-sense psychological standard I will use to assess practicability throughout the remainder of the dissertation.

In Chapter 2, I will outline the ways in which something much like my common-sense standard has been employed in contemporary theorizing about autonomy, contributing to a proliferation of reasonable proposals as to what the exercise of autonomy involves. The additional factors advanced in these proposals have aimed to connect theories of autonomy more closely with lived experience, and have provided the resources through which we can talk about autonomy in a range of difficult, but realistic, cases. I will argue that these elaborations on the nature of

autonomy have come at a cost. Specifically, there were already concerns about the coherence and practicability of autonomy *before* (or without considering) these contributions. Presumably these concerns are exacerbated by the increasingly extensive list of ideas associated with the concept of autonomy. One might worry that increased realism has produced greater incoherence and reduced practicability-related plausibility, thus heightening the need for a system for contending with the seemingly complicated nature of autonomy.

Chapter 3 seeks to provide the kind of system demanded by the concerns about coherence raised in Chapter 2. It is in this chapter that I will elaborate on the schematic provided earlier, offering a method for structuring the analysis of the concept of autonomy that addresses concerns about *extensiveness* by revealing how so many different ideas could be joined together within a single concept.[5] As suggested above, the model I offer has three interconnected levels that provide a way of tracing the relationships between different considerations associated with the concept of autonomy. However, I will end this third chapter by reposing questions and concerns about practicability. While my model helps join different aspects of autonomy together into a coherent picture, it simultaneously reveals the extent to which the exercise of autonomy requires an *extensive* range of abilities and is highly *complex*. Based on the discussion so far, it could appear that from a common-sense perspective we should not expect autonomy to be widely practicable.

Defending autonomy against these renewed concerns about practicability will be the objective of the final chapters of the dissertation. In Chapter 4, I will discuss two features of my conception of autonomy that support the practicability of the exercise of autonomy. First, I remind the reader that I, along with much of contemporary theorizing about autonomy, have adopted the stance that

---

[5] Thus, this is my solution to the second challenge to be faced in theorizing about autonomy.

autonomy is a matter of degree. Thus, we should not be concerned that my view involves the mastery of, or perfection in, the skills and attitudes that are the means by which autonomy is exercised. Only some satisfactory level of achievement is required. Second, I stress how common mechanisms and strategies for reducing the cognitive burden of mental processes such as deliberation and introspection can be employed to facilitate the exercise of autonomy. In Chapter 5, I will introduce a third feature that supports practicability. I will show how people's flaws and weaknesses in exercising autonomy can be overcome by leveraging extra-personal resources to reinforce or substitute for a wide range of deficits. In both chapters, I will endeavour to convince the reader that my proposed features are plausible on a common-sense understanding of human psychology. In Chapter 5 I will conclude with a discussion of what I take the broader implications of my view to be, and I will respond to several outstanding concerns or objections.

In addition to solving problems relating to coherence and practicability in autonomy theorizing, my view makes several contributions that are worthy of note. These include:

- Providing an account of how and why thresholds for qualifying as autonomous vary based on contextual features, which to date is an area that is under-theorized.

- Suggesting a novel form of constitutive relationality that is based on what goes into the exercise of autonomy, rather than on the necessary and sufficient features of the concept.

And, perhaps more speculatively:

- My three-level analysis provides a framework for thinking about autonomy that can be shared by those who see good or right reasons as discovered versus those who see good or right reasons as chosen. That is, whether we see good reasons as objective/external/real or subjective/internal/constructed, the exercise of autonomy in

both cases will involve the use of similar capacities and the exercise of similar

functions. So, I will invite the reader to consider how the use of this framework could

be applied by both procedural and at least somewhat substantive accounts of

autonomy.[6]

- The final picture of autonomy that I have in mind provides the grounds for deeper

  understanding of what it means for autonomy to be a matter of degree and decided by

  different thresholds in different contexts, which I think has implications for our social

  and political responses in instances where people are deemed not to meet particular

  autonomy thresholds.

This list is for now a promissory note and remains to be substantiated. I will begin building

towards these developments, as described above, by addressing questions associated with the

fourth kind of challenge to be faced in theorizing about autonomy: practicability. The next

chapter will take on questions relating to why practicability should be a concern in developing an

acceptable understanding of autonomy. That is, why should practicability be a condition for

plausibility? And, what standard of practicability should our notion of autonomy be required to

meet?

---

[6] Any tendency toward defaulting to constructivist language is a product of my personal stance on these issues, but is not a necessary feature of the three-level model I am advancing. Because the exercise of autonomy may look roughly the same on each view (except perhaps for more or less emphasis placed on certain functions), I do not tend to stress the distinction between substantive and procedural autonomy. My own view would, in the terms of the existing debate, be counted as a weak substantive view because I accept that some positive self-regarding attitudes contribute to our ability to engage in the procedures for choosing authentically and living accordingly.

Chapter 1: Practicability and Autonomy: A First Pass

As discussed in my introductory remarks this dissertation develops a conception of autonomy that aims to be both coherent and practicable. This chapter speaks to the second of these aims. In it, I intend to achieve two principal objectives. First, I will defend the idea that an acceptable conception of autonomy must be informed by empirical considerations relating to practicability. That is, an important consideration in assessing the viability of a conception of autonomy is whether it is not just possible (metaphysically and otherwise), but plausible: that people can qualify as autonomous. Without it, autonomy could only be achieved by some small, exceptional minority of people. Second, I will advance the empirical standard that I will employ through the rest of the dissertation both as a basis for critiquing other views, and as a central criterion in establishing the merits of my own view.

The chapter will proceed in three major stages. In the first section, I will focus on defending the relevance of empirical considerations to the development of an acceptable conception of autonomy. Some will reject this, calling for autonomy to be defined first in the abstract, without regard to practical considerations. However, I will argue that the close relationship between autonomy and psychological concepts, as well as the kinds of thresholds involved in making judgements about whether someone is autonomous, provide good reason to take empirical considerations about human abilities into account. In the second section I will consider two objections to the plausibility of autonomy. The treatment of the first, which is a more historical concern about the metaphysical possibility of autonomy, will be cursory since I take Harry Frankfurt to have largely resolved it in his seminal works on autonomy. The second is an updated version of the kind of behaviourist objection Gerald Dworkin anticipated in "The Concept of Autonomy" (Dworkin 1989, 55). Given the prominence of this second objection in

contemporary discourse about autonomy (and other areas of theoretical ethics), I will offer an extended discussion of the faulty bases on which the view from which this objection originates, situationism, has been developed. I think these flaws severely undermine the credibility of situationists' objections. I will spend the third and final section of this chapter articulating the kind of common-sense empirical standard of practicability that I will use to assess the plausibility of my and others' accounts of autonomy throughout the remaining chapters.

Theorizing about Autonomy: Conceptual and Empirical

Some would argue, much in the vein of traditional conceptual analysis (Bermúdez 2005, 6), that theories of autonomy should be developed purely conceptually. They would hold that autonomy has a nature that is knowable without reference to facts about human nature or abilities, and that we should come to a definitive articulation of the concept prior to considering pragmatic questions such as practicability (Piper 2016, 788). I suspect this kind of approach to understanding the process of formulating a conception of autonomy is attractive to many philosophers, but I fear that this attraction would be misplaced. As I suggested above there are, as I see it, two major reasons why empirical considerations about people and their abilities must be taken into account in the development of an acceptable conception of autonomy. The first, is that autonomy is closely related to psychological concepts. As we will see in later chapters, the skills that are involved in the exercise of autonomy are largely psychological (Heal 2012, 23), and according to José Louis Bermúdez a complete understanding of such concepts can only be arrived at by a combination of conceptual analysis and empirical observation (Bermúdez 2005, 10-11). The second major reason I do not think that a satisfactory analysis of autonomy can succeed without giving due consideration to practicability that autonomy is bound intimately to "the social practices in which statuses, entitlements, immunities, liberties, and the like are attributed..." (Anderson 2014, 355). I will explain each of these reasons in turn.

In his efforts to provide a framework for understanding the relationship between philosophy of psychology and the empirical discipline of psychology, Bermúdez distances philosophy of psychology from approaches to philosophical analysis that claim to be "*a priori*" and aim primarily to specify necessary and sufficient conditions for given concepts (Bermúdez 2005, 6-7). According to Burmúdez, many concepts are neither of the "stipulative and criterial character" associated with this kind of analysis, nor of a perfectly empirical nature (Bermúdez 2005, 9). Psychological concepts are of this hybrid sort, which Burmúdez calls *theory-cluster* concepts. He contends theory-cluster concepts are best understood by considering their roles in the various theories in which they feature (Bermúdez 2005, 10). The skills that are part of the lowest level of analysis that I presented earlier in the schematic outline of my view, are for the most part psychological skills like reason, introspection, and self-control. As such, autonomy appears to "straddle the boundary between philosophy and psychology" (Bermúdez 2005, 9) by virtue of the fact that important components of the concept do so. Even if we do not want to say that autonomy itself is a psychological concept, psychological concepts are importantly implicated in it. This strongly suggests that our conception of autonomy is answerable to what we know, or what we come to know, empirically (Bermúdez 2005, 13). And this, I would contend, means that our analysis of autonomy should admit considerations of practicability based on what we take to be empirically the case regarding people's abilities.

The second reason that I think that it is important for theories of autonomy to take practicability into consideration relates to how autonomy is connected to social practices. Understanding who qualifies as autonomous and how is important because of autonomy's connection to self-realization and to at least certain forms of social status. With respect to the former, if self-realization is part of what it means to live a good life, then we may be interested to know who is

able to achieve this kind of good. For example, this question could be of interest to those who would want to assist those who face barriers to self-realization. With respect to the latter, an important aspect of many political theories is how they define the public space which is left open for autonomous peoples' pursuit of their aims (Christman 2004, 147). Within some constraints (the safety of others, the greater good) the pursuits of autonomous agents are respected and protected from interference by public institutions. Non-autonomous actors are not necessarily granted the same kind of respect, and – for better or worse-- government intervention for the good of non-autonomous individuals is often thought to be justified (Christman 2014, 372). A well-articulated conception of autonomy is therefore needed in order to ensure that our judgements about who is capable of self-realization and who is owed the particular kind of respect associated with autonomy are justified.[7] We therefore need an account of personal autonomy that people can actually achieve, and which does not exclude people for unjust, or otherwise indefensible reasons.

Some would object to this second reason for being concerned about the practicability of exercising autonomy, which defends a concern with practicability on social-political grounds. A recent articulation of the kind of objection against motivating concerns about practicability on the basis of desired social ends has been offered by Mark Piper. Piper rejects the development of conceptions of autonomy based on the idea "that autonomy is fairly common and widespread" (Piper 2016, 788) on several grounds. Although I will not address all of his arguments here, the concern he raises that I wish to respond to at this stage is that defending a low autonomy threshold in order to ensure that autonomy is practicable amounts to smuggling moral or political

---

[7] I am assuming that at least some of our judgements of this nature are justified and that while some of our current judgements about such matters maybe erroneous, they could be justified under an appropriately developed conception of autonomy.

commitments into one's view (Piper 2016, 788). That is, he contends that people who are concerned with the practicability of the conceptions of autonomy they develop are sneaking in what he views as irrelevant factors in order to ensure that their preferred conclusions about autonomy are supported. In Chapter 4 I will reject a minimally demanding conception of autonomy, motivated explicitly by a desire to advance social justice. To a point, my rejection of this position – in agreement with Piper—is based on the belief that normative considerations, such as solidarity, do not on their own justify taking the stance that everyone has the potential to be autonomous. However, as my comments about the close connection between psychological concepts and autonomy suggests, I also disagree with Piper that we can first define autonomy and then determine whether it is widely practicable (Piper 2016, 788). And, this has a special kind of implications for the acceptability of the inclusion of normative considerations in defining autonomy.

While Piper accepts that autonomy is a matter of degree, I do not think that he fully appreciates the implications of this stance. When we conceive of autonomy as a matter of degree, we accept the possibility that across people there will be a range of levels of ability when it comes to exercising autonomy. When thinking about autonomy as a matter of degree, however, it also seems to me that there is no objective level at which a threshold can be drawn in order to distinguish between who is autonomous and who is not. Adopting the idea of autonomy as a matter of degree allows us to understand how, the mere possession (or lack) of certain capacities related to autonomy does not automatically warrant claims about social status. Momentarily, I will argue that what level of ability counts as autonomy will depend on circumstances. Because many circumstances are not very demanding, autonomy is more achievable than recognized by theorists who support a single, relatively demanding threshold. So, the case that autonomy is

likely widely enjoyed can be made without recourse to, or being motivated by concerns about social justice.

As many critics of autonomy have argued, it cannot and— should not—be the case that all questions about the moral and political status of individuals hinges on whether they are autonomous. There are a range of other plausible grounds for moral considerability, such as being conscious or sentient (Frey 2005, 339), or being bound up in broad social patterns of dependence and interdependence (Kittay 2001, 573), by which people we might not deem autonomous obtain important kinds of moral and political status, such as protection against mistreatment and degradation. As such we need not believe, as critics tend to worry, that autonomy is the basis for all forms of dignity and respect.

Especially if autonomy is a matter of degree, therefore, we cannot simply define it and immediately derive moral conclusions from it in the way that Piper contends. Instead, we need to think carefully about the threshold, or thresholds, along a continuum of capability at which people are to be recognized as performing sufficiently well to warrant the particular kind of respect associated with autonomy. It is my contention that there is good reason to believe that we should consider a multiplicity of thresholds since, as Tom Beauchamp says, they are to be "fixed in light of specific objectives of decision making" (Beauchamp 2005, 317). By this Beauchamp means to refer at least to the kind of question a person is considering, which can be as diverse as making medical decisions involving risk, making important financial decisions, or planning for the future (Beauchamp 2005, 317), but also (and on his view more importantly) "everyday choices such as opening bank accounts, purchasing goods in stores, and authorizing an automobile to be repaired" (Beauchamp 2005, 313).

Why the difference between the kinds of instances in which we might be called upon to exercise autonomy provides justification for believing that there are a multitude of autonomy thresholds, however, warrants further elaboration. We might first think of what the consequence of having a single threshold would be. As Beauchamp notes, if the threshold is set at a high level, then it will appear as though many of our day-to-day activities are not autonomous (2005, 322). It will appear that almost no one can function sufficiently well to autonomously open a bank account. Conversely, if the threshold is set at an extremely low threshold, then almost anything will count as an adequate exercise of autonomy. In such a case, we might be led to conclude that children or coma patients[8] can autonomously open bank accounts. The latter case should be intuitively unacceptable, deviating significantly from what we have in mind when thinking about autonomy. The unacceptability of the former case is grounded in the belief, which I share with Beauchamp, that theories that result in deeming the "everyday choices of ordinary people" un-autonomous are dubious (Beauchamp 2005, 312-3). I think, however, that we must go beyond Beauchamp's call to protect everyday choices in articulating why a pluralist understanding of autonomy thresholds is the most appropriate and defensible.

In my view, what counts as a sufficient exercise of autonomy should be based on the nature of the (often social) purposes or practices in question. Specifically, what it means to meaningfully engage in these practices, which relate (among other things) to decision making, self-advocacy, and collective deliberation, will dictate where an appropriate threshold will lie. On my view, contra Piper, practicability remains a concern since if people in general do not perform the

---

[8] Those willing to concede that some non-human animals are able to perform at least some of the capacities associated with autonomy to some degree, might also be led to conclude that even pigs or monkeys can autonomously open bank accounts if the threshold is set sufficiently low, since it is bound to be the case that some non-human animals are more able to exercise these capacities as some humans (Beauchamp 2005, 322; Frey 2005, 341).

functions associated with autonomy to a sufficient degree to participate meaningfully in basic social practices, then it is unclear that our social expectations and practices are properly formulated. It would seem that they exist for other, more perfect creatures, and not for us.

An important first step in developing this position involves specifying what I mean by *meaningful*. In setting thresholds based on meaningful engagement in certain kinds of practices, we are distinguishing between those who succeed and those who fail in this kind of engagement. There are two ways in which someone might fail to meaningfully engage in particular social practices. In my view thresholds are only to be based on the second kind, but I will articulate both so that my meaning is clear.

The first way in which someone can fail to meaningfully engage in some practice is because they experience some form of limitation and, though they may have the social or legal freedom to participate in the practice, they cannot participate because they require unavailable kinds of accommodation or support in order to do so (Knight 2015, 101). This is the difference between having formal versus substantive rights, freedoms, abilities— having the permission or liberty to do something versus having the actual opportunity to do it. One way that participation in social practices has been denied, for example, to people with disabilities is through the absence of the conditions and structures that would make participation feasible to those who fell too far outside the constraints of "majority social arrangements" (Nussbaum 2009, 345).

Critics of autonomy and those who theorize about disability will have a pre-emptive concern that meaningful engagement in the kinds of practices I have in mind will entail an expectation that this is achieved *independently*. However, the discussion in Chapters 4 and 5 will reveal that I hold no such commitment. Instead, I am aiming to show how the exercise of autonomy, in the form of deciding or discovering for oneself what is important and living accordingly, can be

achieved through dependence on supports of various kinds. Thus, failing to meaningfully engage in practices in which one exercises autonomy because of a lack of social arrangements to facilitate alternative ways of undertaking this exercise are not useful for delineating the thresholds of autonomy because they are largely the kinds of institutional failures through which disability is constructed (Knight 2015, 99).

However, there is a second way in which people can fail to meaningfully engage in practices that I think is relevant to establishing autonomy thresholds. This has to do with what conceptually these practices seem to involve. Consider here Jane Heal's account of sporting practices. She notes that these practices are characterized by norms, and might be accompanied by stakes and inducements that help motivate people to abide by the norms (Heal 2012, 24). But, in addition to this, those who are going to successfully engage in these practices will need "complex structures of cognition and habit", and Heal concedes that "[f]ailure or maladjustment" in the required psychological capacities "will make it difficult if not impossible for a person to engage in sporting contests" (Heal 2012, 24).

The practices I would like to consider for the purpose of gaining clarity about the nature of autonomy thresholds range from everyday decision making (Lotan and Ells 2010, 114), to self-advocacy (Redley and Weinberg 2007, 768), to collective deliberation, and even to formal social practices in which we might be construed as exercising our judgement about how best to live in more abstract ways, such as voting.[9] These will share some of the same features of practices outlined by Heal. They will often involve norms, they will have stakes, and there will be outcomes to which people engaging meaningfully in the practice need to be attuned. As a result,

---

[9] If we take voting to be a practice of expressing one's authentic views about what candidate or party has the most compelling vision for their community, etc. If we characterize voting in other terms, like representation of interests, I concede autonomy may not even be relevant.

in order to meaningfully engage in these kinds of practices, people must meet certain levels of functioning with respect to a wide range of cognitive, affective, and other abilities. However, because these practices vary widely in terms of (i) the contexts in which they are relevant, (ii) their stakes and significance, (iii) what we take people to be doing in the course of these practices, and presumably many other factors, different levels of functioning will count as autonomy across cases.

For example, in medical decision making the bare minimum we tend to seek is that people have an "appropriate [level of] understanding" of the question under consideration including a reasonable appreciation of the possible outcomes of different courses of action (Beauchamp 2005, 314). To more fully capture all the dimensions of autonomy, we might also want to say that decision making also requires "the capacity to espouse values and convictions" (Jaworska 2005, 127). However, what counts as an appropriate level of understanding and appreciation of the outcomes, etc. will be far lower in the case of ordering ice cream than in the case of, for instance, making a compelling appeal for physician assisted death. The other practices I have in mind will involve similar, though possibly more, abilities to understand one's situation and possibly the perspectives of others to different degrees depending on the topic and context.

I have argued that the threshold for what level of proficiency counts as autonomy in different cases should therefore be decided with reference to contextual features, and that it will also be guided by social understandings of the practices in question and social expectations surrounding these practices.[10] We need not, in saying this, take our existing practices as the standard for setting autonomy thresholds. It could easily be objected that the norms around our existing social

---

[10] I.e. Normative considerations.

practices are misguided, and that in allowing for these norms to inform autonomy thresholds I am being insufficiently critical. There is a tension, as Cheshire Calhoun writes, in the context of non-ideal moral theorizing between giving due weight to actual practices which may be flawed, and articulating ideal but socially incomprehensible norms (Calhoun 2015, 18). This applies to the way of establishing thresholds I am advancing precisely because the compellingness and very intelligibility of my account relies on an appeal to a shared moral practice. And yet, in so doing I risk importing principles that may be expected to be unacceptable from the perspective of a more idealized hypothetical moral theory.

Those suspicious of autonomy and those attempting to understand the social place of people with disabilities, especially severe cognitive disabilities, will direct criticism at autonomy's alleged connection with ideals such as independence and self-sufficiency. They would argue that valorization of these ideals has pushed efforts to include disabled people in society towards models that emphasise voice, empowerment, self-determination (Redley and Weinberg 2007, 768, Kittay 2001, 560, Knight 2015, 101) and that aim for inclusion by "minimi[zing] the consequences of people's impairments" (Redley and Weinberg 2007, 780). Their critique will continue by pointing out that to varying degrees, in many cases, this approach to social inclusion is inappropriate because some disabilities are a barrier to engaging in the activities prescribed as the means for obtaining social recognition and acceptance. So, in proposing that autonomy thresholds will be set in part by the nature of social practice, I recognize the need to exercise care in selecting which kind of social practices we should conceive of as informing what the thresholds should be. In mentioning above that my conception of autonomy is entirely consistent with dependence and support, I intend to affirm the idea that many people with physical impairments are capable of autonomy. People with only physical impairments and those with

milder forms of cognitive impairment may well be able to participate in liberal forms of social belonging with appropriate support and accommodations. But, I concede, with its strong ties to psychological capacities, some people with severe cognitive limitations may never exercise sufficient levels of autonomy to secure a relative guarantee of non-interference in the pursuit of their wishes.[11]

In light of this, a sceptic about autonomy and an advocate for inclusion would ask what importance autonomy should be taken to have, arguing perhaps that it simply "normali[zes] conventions of the able" (Redley and Weinberg 2007, 780). I appreciate the weight of these concerns; they have informed my position that autonomy is certainly not the ground for moral considerability as a whole, and that concerns about autonomy do not constitute the entirety of morality or justice. I would continue to contend that so long as we think there is something good about those who *can* value pursuing those values in their lives, and so long as we think self-advocacy and collective deliberation in some form are important enough to survive into our considered ideals about social life, then practices involving autonomy (e.g. self-advocacy or collective deliberation) will appear in some form in our idealized accounts of morality. And if, as Eva Kittay acknowledges, stretching these concepts to include the severely mentally disabled "empties them of meaning" (Kittay 2001, 568), there will be some norms that characterize these practices, which someone can fail to meet. Though this is regrettable, it is not nearly as profound a problem as many critics would contend. This is both because (I would contend) autonomy is the deciding factor in only a subset of moral and political questions, and because in framing autonomy as a matter of degree it is plausible that respect for autonomy means we should be

---

[11] It's worth noting that this may sometimes be because they cannot rightly be said to have wishes. Interests, certainly, but these are judged from an external perspective.

sensitive to the degree of autonomy someone *can* exercise, even if this falls below the degree that would count as sufficient in a particular instance.

In this section, I have given an account of why we cannot simply engage in abstract theorizing in order to arrive at a satisfactory conception of autonomy. The first reason is because autonomy is closely linked to empirical claims about human psychology, and needs to be developed accordingly. The second reason is that in treating autonomy as a matter of degree, we need to attend to the kinds of norms surrounding practical judgements about autonomy we are going to make in different contexts, since a single threshold is not going to yield a satisfactory theory. I have argued there are important practical and contextual considerations that will go into establishing these thresholds, and so autonomy theorizing cannot proceed purely on a conceptual basis. This discussion of the nature of autonomy thresholds constitutes the beginning of my attempt to articulate a practicable account of autonomy in which it is compatible with support. My hope is that such an account will be appropriately inclusive. I think however, that it is possible to defend such a stance without it collapsing into an extremely minimally demanding conception of autonomy motivated by imperatives associated with non-discrimination or solidarity.[12]

Together, the connections between autonomy and psychological concepts, and between autonomy and social practices, speak to the importance of being concerned with practicability in articulating a conception of autonomy. These connections not only support Gerald Dworkin's claim that we should reject as inadequate theories of autonomy on which "it [is] impossible or extremely unlikely that anybody ever has been, or could be, autonomous" (Dworkin 1989, 55),

---

[12] The key difference is, I think, suggested by Tom Beauchamp when he says that moral notions should *constrain*, but not fully determine, "how we construct theories of autonomous action and the autonomous person" (Beauchamp, 310).

but also suggests that on an adequate conception of autonomy it may be common for people to function sufficiently well to be considered autonomous in a wide variety of social contexts. That is, I wish to go further than Dworkin's possibility criterion since I would contend that a conception of autonomy would also be unsatisfactory if it allowed that only a very few, special individuals ever exercise autonomy sufficiently well to qualify as autonomous. If the criterion is merely concerned with possibility then it would seemingly be satisfied if only one person is ever autonomous. Expanding the criterion towards practicability would account for our intuition— which by the end of the chapter we might claim is part of our common-sense psychological framework— that people, with the exception of infants and those with significant mental limitations caused by (for example) senility, severe developmental delays, insanity, or brain death, "*are*, for the most part, autonomous" (Nahmias 2007, 169, Feinberg 1989, 28).[13] On such a conception, autonomy is not merely possible for beings like us, but can be plausibly ascribed to many of us, much of the time. The next section will consider and counter challenges mounted against the possibility and plausibility of wide-spread autonomy.

Metaphysical and Social Psychological Objections to Autonomy

It appears that in proposing the criterion of empirical possibility as a measure of an adequate theory of autonomy Dworkin had at least two potential objections regarding possibility in mind. One of these objections was about the need for people to be free from the effects of socialization to count as autonomous —which is deeply inconsistent with lived experience (Dworkin 1989, 55). Such views, he contends, hold that being autonomous involves being an "unchosen chooser, [an] uninfluenced influencer" (Dworkin 1989, 58). This constitutes a problematic conception of

---

[13] Granted, the conditions under which such a claim is interpreted as true might differ depending on the conception of autonomy being employed. For example, my view holds that autonomy depends on adequate functioning and so my interpretation of what it takes for people to be for the most part autonomous will be significantly different than an interpretation based on a conception that holds autonomy to be a form of mastery.

autonomy because it relies on a kind of metaphysical freedom from causal forces. Second, Dworkin was aware that certain experimental paradigms, such as B. F. Skinner's behaviourism, deny the existence of autonomy altogether on empirical grounds. This is a source of doubt that Dworkin takes quite seriously, and so he advocates that experimental paradigms and results should be taken into account when assessing whether a particular conception of autonomy is acceptable (Dworkin 1989, 55). While these two objections to autonomy are quite different, it is worth noting that they are united in their interest in the potential problems that causality poses for autonomy.

I will treat the metaphysical problem, which I take to truly doubt the sheer possibility of autonomy, relatively briefly because in "Freedom of the Will and the Concept of the Person" Harry Frankfurt provided the necessary distinction between autonomy and metaphysical freedom that remains in use in subsequent contemporary theorizing about autonomy. My main goal in this section is instead to address an updated and very live version of the concern Dworkin was entertaining related to behaviourism. In recent years situationists like John Doris and Gilbert Harman have made much of empirical findings that appear to suggest that peoples' personal values do not cause their behaviour in the way that often seems to be assumed in discussions about character. As I will explain, this position has implications for the practicability, or even outright existence of, autonomy. While stronger versions of the situationist position would contend that autonomy does not exist, there are also more moderate views that maintain that the scope for autonomous action is considerably more restricted than we might assume. The truth of either position would constitute a considerable blow to the plausibility of autonomy. Although Dworkin did not frame his criterion of adequacy in terms of plausibility, the move from mere possibility to plausibility is justified not only due to autonomy's connection to psychological

concepts and social practices, but also because some situationists are willing to accept that autonomy may be possible, just not widely enjoyed. I will argue, however, that the evidence that situationists rely on to establish their views is not sufficient to raise genuine doubts about autonomy.

i.   A Metaphysical Objection to Autonomy

The metaphysical concern about the plausibility of autonomy as a state or power that people actually enjoy arises because of the connection between contemporary accounts of personal autonomy and philosophical debates about freedom of the will (Christman 1989, 7). Indeed, the seminal papers by Harry Frankfurt were composed as explanations of what it means to have freedom of the will, though they also shed light on the nature of personal autonomy (see for example Frankfurt 1971). If we think that autonomy is the same as free will, then autonomy will be under the same suspicion as accounts of free will from a metaphysical perspective. That is, there is a tradition of claiming that free will involves contracausal freedom, or "the absence of causal determination" (Frankfurt 1971, 18). It contends that to have free will agents will act at least partly of their own accord in a way that is not merely a product of the chain of causes that have brought them to a particular moment of action (Stace 1953). Frankfurt attributes such a view to Roderick Chisholm (Frankfurt 1971, 18) and W.T. Stace claims this view was a common assumption throughout the modern period (1953). The metaphysical objection to this view of free will is that it appears to be metaphysically impossible, requiring that free agents have God-like powers or the ability to step outside the operation of natural laws (see for example, Berlin 2002, 184). And, according to the objection, if free will is impossible then we cannot be autonomous (Nahmias 2007, 169).

Frankfurt's view of the relationship between free will and responsibility shows why autonomy is not open to this metaphysical objection in the way that freedom of the will might be. For Frankfurt, exercising freedom of the will involves aligning the desires on which we act (our effective desires) with our second-order volitions (the desires about what kinds of desires we act on) (Frankfurt 1971, 15). For example, many drug addicts, according to Frankfurt, do not have freedom of the will in that they will act according to their drug-seeking desires regardless of whether they want, on a higher order, to have these desires at all (Frankfurt 1971, 15 & 19). *However*, Frankfurt notes that one can lack freedom of the will and yet still be morally responsible. Returning to his example about drug addicts, Frankfurt argues that an addict who desires that their first-order drug-seeking desires are effective takes ownership of their drug-seeking desires in a way that an addict who repudiates their drug-seeking desires does not; these desires are alien to the unwilling addict. Through the ownership established by a second order volition, Frankfurt's "willing addicts" come to be responsible for their drug-seeking desires and behaviours (Frankfurt 1971, 20). On my reading[14], responsibility is a synonym for, or overlaps with, autonomy in this part of Frankfurt's discussion. Frankfurt's account suggests that while autonomy may share similarities with freedom of the will, the two can come apart, insulating autonomy from problems relating to determinism and freedom of the will.

Although there are some flaws with Frankfurt's presentation from the perspective of autonomy, for example we might come to worry about the authenticity of a willing addict's willingness, it is a useful illustration of how autonomy is not made impossible or even implausible by the fact that autonomous agents will be influenced by causal laws or that their actions are determined. The

---

[14] This seems like Dworkin's reading as well, though he doesn't refer to "Freedom of the Will and the Concept of a Person" directly. See Dworkin 1976, p. 25. Arpaly's critique of the synonymy of autonomy and moral responsibility will be discussed in the next chapter.

next objection to autonomy I will consider questions whether, having endorsed certain desires, beliefs, values, or goals—which will be referred to collectively as "self-defining attitudes" following Suzy Killmister (Killmister 2018, 14)— we can reasonably expect they will shape our behaviour. I will be responding to a situationist objection of this nature, which contends that empirical research into human psychology provides evidence that our ability to shape our behaviour in accordance with our values is limited, if it exists at all. This would be cause to doubt that people can live according to their authentic values, which is part of autonomy.

ii. *Social Psychological Objections to Autonomy*

When Dworkin outlined his empirical possibility criterion, the objection to autonomy originating in psychological research he had in mind was informed by the work of B.F. Skinner (Dworkin 1989, 55), whose behaviourist paradigm proposed that all causes of behaviour are external to the agent (Ryan and Deci 2006, 1568). On Skinner's view, behaviour is primarily shaped by the environment in which we act, rather than by the principles to which we are committed (Nahmias 2007, 169). If this were the case, then it would seem we do not possess the ability to be self-realizing in the way that contemporary views of autonomy describe. Eddy Nahmias notes that the Freudian psychological picture also raises questions about how autonomous people are, since self-governance does not seem compatible with being driven significantly by unconscious desires (Nahmias 2007, 169). If our reasons for acting are not what we take them to be and instead we are driven by sub-conscious processes, then it seems hard to believe that we are able to purposefully live in accordance with our values since they are not the basis for our behaviour. Although "the influence of Freud and Skinner has waned" (Nahmias 2007, 169), similar objections that our behaviour is mostly the product of environmental circumstances and that it is importantly the product of sub-conscious (or pre-conscious) processes can be made based on

more recent social psychological research. John Doris and Gilbert Harman have been important figures in drawing out the implications of this research in philosophical discourse as part of advancing their situationist critique of the place of character in moral discourse, especially in relation to virtue ethics (Doris 1998, 506, Harman 1999, 316).

Doris, Harman, and other situationists cite a number of studies that appear to disconfirm the existence of character traits as they are commonly conceived of. They use these empirical findings to support their arguments against moral psychological accounts that postulate the existence of character traits (Doris 1998, 505). On their interpretation, these findings support three counterintuitive conclusions about the role of character in behaviour (Harman 1999, 319). First, people's behaviour does not appear to be governed by their traits. Behaviour is more reliably predicted by aspects of their situation (Doris 1998, 507). Second, people who are attributed a certain trait do not behave reliably in accordance with this trait (Doris 1998, 508). And third, people may have both the virtuous and vicious version of a particular disposition on a more fine-grained analysis of their character (they may be honest in some circumstances and dishonest in others, for example) (Doris 1998, 509). Situationists therefore argue that behaviour is primarily a product of environment rather than of character (Doris 1998, 508). This leads proponents of situationism to conclude either that "all thought and talk of character and virtue" should be abandoned (Harman 2000, 224), or at least that characterological terminology should only be applied to "actions, institutions, …states of affairs", and fine-grained dispositions[15] (Doris 1998, 514).

---

[15] On this account, a person might have a character trait characterized as courage-in-speaking-in-front-of-the-class, rather than a global trait of courage.

Although John Doris claims that situationism does not reject the idea that people are "autonomous agents who can, in some deep sense, chart the course of our own lives" (Doris 1998, 515), it is hard to see how this could be the case. The findings situationists read as undermining the existence of character traits also bring the plausibility of autonomy into question. The relationship between autonomy and character can be reconstructed by attending to some of the traits or powers that we would reasonably expect self-governing people to have. If autonomy involves authentically choosing or endorsing values and living according to them, then we must expect that an autonomous person can act according to their principles. That is, an autonomous person's self-defining attitudes should guide her actions (Nahmias 2007, 170), especially those actions that she takes as involving the deliberate exercise of her autonomy. The literature that situationists cite in raising doubts about character bring this aspect of autonomy into question in several ways.

First, the empirical results that suggest that situational factors are a stronger determinant of behaviour than character undermine the idea that we can "act on consistent principles we endorse" (Nahmias 2007, 175). When the strongest shared attribute among people who display the same behaviour is whether they think they are in a rush, or whether they recently found some change, it's hard to believe that people's actions are directed by their values as our view of autonomy would expect. Doris's claim about the compatibility of situationism and autonomy involves placing more emphasis on using our values to choose which situations we end up in so that we will be caused to act in the ways we endorse. He urges we should avoid situations in which we will be caused to act contrary to our values, and seek those in which we will do good

(Doris 1998, 517). While I am in some ways sympathetic to this idea[16], I question the satisfactoriness of the kind of autonomy we would have on his account. The best we would seem to be able to do is attempt to foresee and avoid circumstances in which we are likely to act contrary to our values. Autonomy of this kind would be dependent on our ability to accurately make predictions about the future, and it would be close to impossible to exercise autonomy in substantially new or unexpected circumstances. Autonomy understood in this way would therefore be considerably less robust than is suggested by our ordinary usage.[17]

The second way in which the situationist model undermines autonomy is in pointing to research that shows people behaving regularly so long as the situation remains "substantially similar" (Doris 1998, 507, Nahmias 2007, 175). While the similarity in our reactions in these situations might be guided by the same beliefs or values, we don't tend to think that the values we espouse are highly situation-specific. We tend to think that these values are supposed to guide us in a wide variety of contexts (Nahmias 2007, 175). But, according to situationists, the empirical evidence does not support the idea that our values structure our lives and our behaviour in these general ways.

A final problem that the empirical findings cited by situationists create for the plausibility of autonomy is that they indicate that largely people are unaware of the situational factors that guide their actions. This suggests that our behaviour and our decisions may be the result of subconscious or unconscious processes, which would mean that whether we act in accordance with our principles could be a matter of chance, or that our explanation of our behaviour in terms

---

[16] My own account will employ a version of modifying one's environment (situation) in order to exercise autonomy, to be discussed in Ch. 5.

[17] Situationists might be quite satisfied with this conclusion. But, given the nature of this project it should be clear that I do not agree that this is a helpful development.

of principles involves *post hoc* rationalization (Nahmias 2007, 174). Indeed, the inaccuracy of the subjective reports people make about their reasons for having acted (Nahmias 2007, 176) might raise doubts not only about our ability to live according to our values, but also about our ability to select our values authentically since this at times requires introspection about the nature and origins of our values and the ability to critically reflect on them. These functions would also rely on reliable introspection, which situationists' empirical sources do not support.

More conciliatory situationist critiques of character suggest that even if autonomy is not entirely discredited as a human ability, our autonomy is at risk or threatened to the extent that empirical observations reveal "limitations on our capacities to form consistent principles" and to act on them (Nahmias 2007, 175). But, this weaker version of the objection still constitutes a significant challenge to the expectation that autonomy is widely enjoyed. Given that my aim in the following chapters is to develop a practicable and productive conception of autonomy for use in philosophical and interdisciplinary contexts, I am obliged to answer the situationist challenge. There are, as I see it three approaches to this task. The first approach, which would be to question the general relevance of psychological findings to philosophical analysis, is closed because I have already admitted the close relationship between autonomy and psychological concepts, as well as made a commitment to developing an empirically plausible account. So, I must have an interest in what empirical observation has to say. And besides, the work of Doris and others may have already shaped the philosophical landscape in such a way as to make "purely conceptual argumentation and armchair speculation" unacceptable in moral philosophy (Snow 2013, 340).

The second approach involves scrutinizing the quality of the studies cited by situationists. Critics of situationism have pursued this avenue on at least two different levels. The first level of

critique has to do with methodology. An example of a critique of this kind is how Nancy Snow has criticized Harman and Doris's use of a 1972 experiment by Isen and Levin. Her complaint about the methodology of the study is that it had a very low number of participants and as such is not very reliable (Snow 2013, 348). Further methodological concerns are raised by Eddy Nahmias. He wonders whether subjects in these studies were set-up to fail to demonstrate character since they were not prompted to attend to what they were doing in advance and were not necessarily made to respond to problems they had already considered or already cared about. Nahmias suspects that subjects under conditions where they knew their behaviour mattered, or who were exposed to a situation they cared about and had already considered, would have been better able to act according to their character (Nahmias 2007, 178).

We also might have methodological concerns about the often cited 1973 experiment with Princeton seminarians by Darley and Batson. This study seems particularly at risk of running afoul of the problem of WEIRD sampling, though other work cited by situationists may be guilty of this as well. Henrich, Heine, and Norenzayan have argued that the common practice of conducting psychological research using western, educated, industrialized, rich, and democratic subjects (usually undergraduates in university psychology courses) to establish psychological universals undermines psychologists' "ability to distinguish…reliably developing aspects of human psychology from more developmentally, culturally, or environmentally contingent aspects of our psychology" (Henrich, Heine and Norenzayan 2010, 61-63). Henrich, Heine and Norenzayan contend that psychology is one of several fields guilty of selecting extremely biased subject pools and then claiming to establish a universal about human nature, much to the detriment of the credibility of the field. Using seminarians at Princeton to investigate the role of character in human behaviour strikes me as a prime example of this. Specifically, I'm suspicious

that the student body of the Princeton seminary in 1973 would have been largely, if not entirely, WEIRD. It would however, be necessary to examine the methodology and data from all the studies cited by situationists to ascertain the full extent to which their sources fall prey to this problem.

A second level on which the studies cited by situationists can be scrutinized relates to whether they actually show what they are purported to.[18] For example, Kingsbury and Dare worry that Harman's claim that the Princeton seminary experiment supports the conclusion that there are no character traits is unjustified. They note that there is room to attribute "the character trait of being compassionate even when under pressure" to at least 10% of the subjects (Kingsbury and Dare 2017, 456). Compassion-when-under-pressure is not nearly as narrow a character trait as situationists like Doris are willing to accept, so it appears to conflict with a situationist interpretation of the results.[19] Nahmias raises further concerns on this level relating to experiments cited by situationists in which subjects were asked to explain their judgements after the intervention, and in which the incongruence between their explanations and the measured causes of their actions were interpreted as a lack of introspection. Nahmias notes that introspection and retrospection may be importantly different, and as such the results may have been different if subjects were asked to give their reasons for acting before or while they acted (Nahmias 2007, 178).

In addition, Saigle et al. raise questions about the interpretation of experiments conducted according to the Libet paradigm. Libet and colleagues showed that people's brains form an

---

[18] Some of this might be a problem with situationists' interpretation of the studies rather than the study authors' own interpretations.

[19] Granted, this interpretation doesn't support the present project much either since it still doesn't entail that autonomy is widespread. But, this point does help convey the idea that more than just the situationist reading of experimental results is possible.

action potential before they consciously choose to act. Situationists have taken this to imply that our personalities are not the ultimate source of our actions. However, Saigle et al. note the significant difference between the kinds of finger moving exercises studied in the Libet experiment and the substantially more involved actions and decisions taken in everyday life. They argue that it is unreasonable to draw strong conclusions about the possibility of voluntary action from the Libet experiment since, even if we accept the implications for conclusions about finger twitching, this does not mean the same applies to more complex forms of behaviour (Saigle, Dubljevic and Racine 2018, 38).

These examples of alternative readings of what was observed during the Princeton seminarian experiment, the Libet experiment, and others illustrate the difficulties and snares that are associated with articulating the significance of empirical results. Kingsbury and Dare's concerns about the interpretation of the Princeton seminary experiment show how the same finding can be interpreted and stated in different ways, opening the risk of imposing one's theoretical orientation on how the findings are construed. The kinds of interpretive problems raised by Nahmias, as well as Saigle and her co-authors, point to the difficulty of establishing what kinds of data are relevant in the first place, and how justified we are in interpreting the findings as applying beyond the conditions in which our observations are made.

These criticisms relating to methodology and interpretation represent just a sampling of the worries that have emerged (or may emerge) from a close inspection of the sources drawn on in the situationist critique of character. It would be possible to further undermine the situationist position by conducting a systematic methodological and interpretive critique of their main sources. I think the more important and forceful argument to be made against situationism, however, has to do with *how* proponents of this view have treated the empirical sources on which

their view is premised, as well as the larger body of psychological research on character—this is the promised third way of responding to the situationist challenge.

The third approach to critiquing the grounds on which situationists rest their arguments relates to the *way* they have drawn on scientific findings. There are at least four problems with the way that science has been used in the situationist literature, undermining the credibility of their critique of character in moral philosophy and moral psychology. The first and most minor of these problems is that situationists in some instances overstate the strength of their evidence. Returning to the problem with the Isen and Levin experiment noted above, while the number of the participants was low (~41 people), Snow notes that Doris, Harman, and Merritt report the outcomes of the study as percentages in the section on character they co-authored in of *The Moral Psychology Handbook*. This makes the findings seem much more significant than they actually are given the limited sample (Snow 2013, 348). One of the ways in which Kingsbury and Dare think that Doris overstates the strength of the evidence is by making unjustified appeals to the authority of science. The studies he cites have not become matters of scientific consensus; they have merely been published reporting their results. According to Kingsbury and Dare, the claim, for example, that "studies have shown" is only justified when the results have undergone the rigorous process of "peer scrutiny and [re-]testing". Before then one is making an appeal to "underdeveloped" science (Kingsbury and Dare 2017, 453). The empirical claims made in arguments advancing the situationist position need to be scrutinized in order to ensure they are accorded only the weight they actually merit.

The second way prominent situationists have been accused of misusing their sources is by cherry-picking their results. I was suspicious of this when I noted the same handful of studies being cited repeatedly in the sampling of articles by Doris and Harman I used to familiarize

myself with the situationist position. My suspicions were later confirmed by Kingsbury and Dare as well as Snow. The former authors note that Doris only attends to confirmatory replication attempts of the 1972 Isen and Levin experiment, ignoring that the efforts to replicate the results have for the most part been unsuccessful (Kingsbury and Dare 2017, 451). Snow levels a similar charge against Harman, who continues to cite the 1973 Darley and Batson experiment, but ignores that *even Batson himself* revised his interpretation of the study (in a way more consistent with the existence of character traits) based on subsequent work (2013, 349). Snow also notes a broader problem which is that the situationist literature "tend[s] to ignore" empirical psychological studies that are consistent with the idea of character traits (2013, 349). So, to see the situationist perspective as convincing we would need to hear how they plan to contend with contradicting empirical findings.[20]

The third and fourth problems with the way empirical research tends to be treated by situationists are interconnected. The third problem, briefly, is that the sources they rely on are not very current. Even in *The Moral Psychology Handbook*, presumably *the* venue in which to present "state of the art essays" (Snow 2013, 341), the five representative experiments placed front and centre by Doris, Harman, and Merritt were *all* conducted in the 1970s (2010, 357-58). While there is no expiry date on empirical research, it would be beneficial for readers to see evidence of progress in this empirical tradition which Doris claims is now over 70 years old (Doris 1998, 504). Without more recent sources readers are left to wonder why, in the last *forty or more years* there are no findings that proponents of the situationist position think are significant enough to tout in their piece that is presumably supposed to represent the state of contemporary discourse.

---

[20] Doris at least seems disinclined to pursue this task having, in his introduction to the Moral Psychology Handbook, eschewed the "fool's errand [of] the pretense of impartiality" (John Doris and The Moral Psychology Research Group 2010, 2).

The empirical findings cited by situationists would be more credible, and this brings me to the fourth problem with the way situationists tend to treat empirical research, if there were many replication studies or meta-analyses based on the studies situationists tend to draw on demonstrating consistent significant effects. Indeed, Kingsbury and Dare note that if "an established pattern of repeatable results" had been found, the situationist case would be much better supported by a report on this rather than by direct reference to studies, like the 1972 Isen and Levin study, they have been re-citing for decades (Kingsbury and Dare 2017, 451). While Doris appears to be sensitive to the importance of replication, and claims to only use studies that have been replicated or extended (Doris 2005, 30 n.3), his use of sources, particularly the Isen and Levin study, has come under scrutiny. The reason for this is that, while this study has been successfully replicated, most replication attempts have been unsuccessful. By the end of the 1970s, Kingsbury and Dare note, psychologists were drawing the conclusion that the failures to replicate were enough to bring the result of the original study into question (Kingsbury and Dare 2017, 451). So, while Doris's awareness of the importance of replication is laudable, the standard to which we hold situationists' evidence needs to be quite a bit higher than what situationists have offered so far.

The empirical grounds of their arguments, based on older, problematic, unreplicated (or unreplicable) studies, should not be accepted. Psychology as a discipline is in the process of coming to terms with a web of interconnected methodological and institutional problems that are reducing psychologists' own confidence in the reliability of reported findings in their field (Maxwell, Lau and Howard 2015, 487). The problems have included publication bias (the tendency of journals to publish only articles that report a statistically significant result) and other perverse incentives connected to how social, financial, and occupational rewards are achieved in

the research community, not to mention questionable research and reporting practices and instances of out-right fraud (Spellman 2015, 889-90). A major concern stemming from the acknowledgment of these factors is that there is "likely [an] overabundance of false positive results in the scientific literature" (Maxwell, Lau and Howard 2015, 487). Even if research was conducted well, it would be more likely to be published if it reported a successful (that is, statistically significant) outcome. Replication is an important part of weeding out these false-positives. However, a single confirmatory (or disconfirmatory) replication is not very informative since its results may also be skewed, possibly even due to chance (Maxwell, Lau and Howard 2015, 488, Earp and Trafimow 2015, 2). As such, Doris's attempt to bolster his claims by attending to studies that have been replicated is not adequate.

Especially given the strength of some situationist claims (like Harman's total denial of the existence of character) what is needed in order to provide a firm empirical basis are meta-analyses (or at least comprehensive consideration) of multiple well-designed attempts to replicate the same findings (Maxwell, Lau and Howard 2015, 495, Earp and Trafimow 2015, 9). It does not seem that such analyses are readily available, and where they are they may not be favourable to situationists. For example, Saigle and her coauthors attempted something like the necessary kind of review for a single experiment frequently cited by situationists, the 1983 paper by Libet in which people were found to have neural readiness potentials to act before having consciously decided to act (Saigle, Dubljevic and Racine 2018, 30). The team found considerable variability among the methods and equipment used in the studies even though they were drawing, usually explicitly, on Libet's methods (Saigle, Dubljevic and Racine 2018, 37).[21] Saigle and her co-authors point out that this variability in how the experiments were carried out makes

---

[21] Or else often a more recent study which itself was a replication of Libet's work.

it difficult to say that the original experiment has been successfully replicated. Their work highlights the difficulties of meeting the standards set by emerging norms in psychology and cognate disciplines, perhaps especially when drawing on research that was not designed with these emerging norms in mind.

In this sub-section, I have presented the problems with premises supporting the situationist stance against autonomy in an effort to dispel concerns that experimental evidence suggests that autonomy does not exist or is extremely limited. I have discussed two major kinds of problems with the literature that situationists tend to cite. The first are problems with data quality, where the studies to which situationists point are methodologically unsound. Some of the weaknesses of the studies include small sample sizes, biased samples, and potentially ineffective prompts to research participants. The second involve problems with interpretation, where it is not clear that we should draw the conclusions situationists have from the experimental results. In addition, I outlined four broader concerns with the improper ways in which situationists have tended to use empirical findings generally, further undermining their position. We saw that in some cases situationists have overstated the strength of the empirical evidence, that their arguments face accusations of cherry picking, that their most cited sources do not show evidence of scientific progress, and that the sources they use do not live up to current standards of reliability that are gaining prominence in psychology and related fields.

Given that situationists rely heavily, one might even say exclusively, on empirically based premises to undercut the plausibility of autonomy, we can only conclude that the situationist case provides insufficient evidence of adequate quality to be given much credence. I acknowledge that the deficiencies in situationists' empirical evidence, particularly in regards to the use of meta-analyses of well-designed replication attempts, are partly innocent in that from what I can

tell, the necessary experimental work and meta-analyses have not yet been conducted.[22] The

Saigle et al. review I mentioned earlier addresses so-called replications of a single experiment

situationists take as evidence that we do not act based on our convictions but are heavily

influenced by unconscious factors. While Saigle and her co-authors point to problems stemming

from the variations in methodology and other aspects of these replications, their findings are also

not sufficient to refute the validity of this approach altogether (Saigle, Dubljevic and Racine

2018, 38). And, even if they were, this is only one of many sources of empirical support on

which situationists base their position.

A full treatment of the sources relied on by situationists would potentially require many meta-

analyses all of which will have to contend with or take stances on further problems relating to

what counts as a replication, which methods are valid[23], and how results are to be interpreted.

But, these have largely yet to be conducted. The seeming absence of empirical evidence of

sufficient quality also has important implications for how I will develop my view moving

forward. Specifically, it raises questions about how judgements about practicability (and

therefore plausibility) will be made throughout the rest of the dissertation.

Adopting a Common-Sense Lens

In my efforts to articulate a coherent and practicable conception of autonomy, I will be referring

to what we might take to be people's, often mental, skills and capacities that are associated with

the exercise of autonomy. One way of establishing that the conception of autonomy I am

---

[22] After submitting this dissertation for defense I did find a meta-analysis of helping behaviour prepared by Lefevor et al. It provides an explanation as to why psychologists have not and likely will not put much effort into establishing the more robust empirical findings that situationists require. The authors are not aware of any psychologists who directly claim that "situational causation" is the primary factor in helping behaviour and suggest that psychologists might simply see philosophers as having misread the literature (Lefevor, et al. 2017, 230). The meta-analysis also found that "situational factors, taken alone, cannot explain helping", which the authors take to make clear the empirical inadequacy of situationism (Lefevor, et al. 2017, 252).

[23] In the sense employed by empirical researchers meaning that they actually measure what they claim to.

developing is practicable would be to seek out psychological literature that speaks to what people's mental capacities are and what their limits tend to be. Yet, the foregoing discussion illustrates the difficulties involved in using empirical evidence to address the kinds of questions to be raised in this dissertation. It is not always clear how to interpret the evidence that does exist, and recent changes in standards of evidence may mean further research or forms of analysis are required to obtain authoritative evidence. So, in what sense can I contend that the resulting framework is empirically plausible? In what remains of this chapter, I will articulate the common-sense approach I intend to take toward assessing and establishing plausibility for the purposes of this dissertation. While it is not uncommon for contemporary autonomy theorists to claim to make use of common-sense[24] in articulating the nature and demands of autonomy, it is worth pausing to explain what precisely I take this to mean in order to avoid giving the impression that I am being overly simplistic about human psychology, or overly credulous regarding the extent to which people can have access to their own mental life.

The understanding of common-sense that I plan to take is reminiscent of the platitude account of folk psychology (Stich and Nichols 2003, 239). The platitude view of common-sense (folk) psychology has its origins in David Lewis's 1972 paper "Psychophysical and Theoretical Identifications", in which he articulates a functionalist view of how theoretical terms come to be defined (Lewis 1972, 249). On Lewis's view, theoretical terms are defined in terms of their "causal (and other) relations to" other terms within the theory as well as terms that exist outside of the theory (Lewis 1972, 253). The most important part of Lewis's view for my purposes is the idea that mental state terms— the ones that are involved in platitudes about mental life that are

---

[24] Explicit reliance on "common-sense" can be seen in other works on autonomy such as Wolf 1989 (p.150), Meyers 2005 (p. 45)

common knowledge (Lewis 1972, 256) or "that almost everyone would agree with and take to be obvious" (Stich and Nichols 2003, 238)— are taken to be theoretical terms, which occupy functional roles in predicting and explaining human behaviour (Lewis 1972, 257-58, Bermúdez 2005, 17). We regularly make use of elements of this often tacit theory "to navigate the social world" (Bermúdez 2005, 10).

Bermúdez notes that there are at least three different ways in which common-sense psychology can be understood among psychologists and philosophers of psychology. The first takes common-sense psychology to refer to the abilities that *underlie* our capacities for social understanding and coordination (Bermúdez 2005, 175). The second takes common-sense psychology to be "a conceptual *framework* for social understanding", while the third depicts it as a method of applying this framework for the purposes of explanation and prediction of behaviour (Bermúdez 2005, 176). The version of the platitude view I am working with is most closely associated with the second way of understanding common-sense psychology. Common-sense psychology (on this platitudinous view) therefore offers a kind of personal level explanation about the functions people can perform, rather than an explanation at sub-personal levels having to do with "cognitive mechanisms" or the physical structures on which the cognitive mechanisms are based (Bermúdez 2005, 17). Personal level explanations are only possible for entities with minds, and roughly speaking, personal level explanations can be distinguished from lower level explanations by their appeal to states that (i) are conscious or at least potentially conscious, or (ii) change in response to changes in people's beliefs and desires, or (iii) cause changes in beliefs and desires (Bermúdez 2005, 30-31). The concepts employed in common-sense psychology are the means by which we describe the intentional domain and the relationships between its parts (Malle 2006, 208). In addition to offering personal level explanations, the other hallmark of

common-sense psychology is that "[i]t picks out classes of behavioural regularities" (Bermúdez 2005, 34). It is through reliance on these regularities that truisms and platitudes about people's motives and behaviours that our common-sense psychological notions aid us in navigating the social world (Bermúdez 2005, 178).

I will be adopting a slightly expanded approach to common-sense psychology. The notion of common-sense psychology described above is behaviourist in origin (Goldman 2000, Jackson and Pettit 1990, 31), as such it is primarily aimed at articulating the concepts we can use to "work backwards in explanation from people's behaviour to their desires and beliefs" (Bermúdez 2005, 33). I would reject that these are the only kinds of personal level states we can know. For example, behaviourist common-sense psychology does not include platitudes relating to (or more generally account for) those mental states we can know through introspection. It will also not include our ordinary beliefs about what mental processes there are and how they relate to each other (Rips and Conrad 1989, 192), and other "assumptions… social perceivers make about human capacities" (Guglielmo, Monroe and Malle 2009, 449) which include not only propositional attitudes, but choices and skills for carrying them out. These further phenomena constitute important aspects of common-sense explanations of people's lives that are neglected by purely behaviourist accounts (Guglielmo, Monroe and Malle 2009, 459). Some of the platitudes that will be relevant for our purposes include ideas along the lines that to some extent people can: exercise self-control, make reasonable judgements about things that affect our lives, and know ourselves and the world well enough to pursue our ends successfully. These are psychological constructs that we have and regularly use that presumably cannot be read off of behaviour.

Adopting a common-sense view does not necessarily mean we can simply appeal to our everyday beliefs uncritically and whole-cloth. Reliance on our common-sense understanding and specification of the extent of our abilities can be supported by or grounded in arguments. We will see this in the next chapter, for example. And, although the account given above supports the usefulness and general reliability of common-sense psychological concepts and judgements, this does not mean that common-sense psychology, both of the conventional behaviourist sort and of the sort I will use in this dissertation, is without limitations or flaws. Here I will consider two limitations of standard common-sense psychology which also extend to my claims about common-sense psychology. These have to do with the relationship between personal level explanations and explanations at lower levels, as well as with whether we think that personal level descriptions suggest that the causes of peoples' actions are always transparent to them. Finally, I will also consider a limitation of my distinctive usage of common-sense psychology that relates to how it departs from more standard conceptions.

First, although common-sense psychology resides at "the top of the hierarchy of psychological explanation" (Bermúdez 2005, 33)[25], different ways of conceiving of the mind adopt different accounts of the relationship between common-sense psychology and sub-personal explanations of human psychology offered by psychological, cognitive, and neuro-scientific approaches. These other approaches specialize in describing the lower levels of the hierarchy (including brain mechanisms and structures). For my purposes, it is not necessary to take a stance on the nature of this relationship. As such, it is important to note that as a result the scope of my claims using common-sense psychology is limited. In particular, making a claim about common-sense

---

[25] I.e. Psychological explanation might be conducted at different levels that are ordered hierarchically. As described by Bermúdez, in this hierarchy personal level explanation is at the top, sub-personal psychological mechanisms in the middle, and brain structures.

psychology does not imply anything about how the brain is structured, or how sub-personal mental mechanisms work. That is, my usage is not concerned with "the 'proper' metaphysical status of mental states" (Goldman 2000). The folk-psychological platitudes I will consider may well be the subjects of dispute in cognitive science or neuroscience. Consequently, it is important to keep in mind that how mental state terms reduce down to, or map on to, for example, neurophysiology, falls outside the scope common-sense psychology of the stripe I am employing (Jackson and Pettit 1990, 38).

Second, adopting a common-sense psychological approach should not be taken as equivalent to suggesting that people have perfect access to their mental states. As we saw with the situationist critique of autonomy, one of the primary bases for their objections to autonomy is that people's behaviour seems to be swayed by factors of which they seem to be entirely unaware. There are at least two senses in which this is true. First, the form of understanding that common-sense psychology involves could be at least in part implicit (which points back to the idea that personal level states can be merely *potentially* conscious[26]) (Bermúdez 2005, 180-81). Thus, we may not have an explicit list of all the trivia that we take to be true about people, but may recognize them as true after the fact or on reflection.

Furthermore, common-sense psychology is entirely consistent with the idea that our actions and motivational states can be the result of causes of which we are not conscious (Malle 2006, 227). There are remain outstanding questions, specifically about how to interpret experimental results, about the nature of the relationship between awareness and the causal role of intentions (Malle 2006, 209-10). Malle, for example, does not think that the fact that people are vulnerable to

---

[26] Though Bermúdez also remarks that implicit knowledge or understanding is not clearly defined in this or connected fields.

unconscious influences means that they are under the control of these influences any more than under the influence of their own conscious intentions (Malle 2006, 213). What implications our inability to access all factors that are involved in our mental lives have for our beliefs about the reliability of common-sense psychological explanation remains the subject of debate. But, I will not take up this debate here. I think at this stage it is sufficient for me to acknowledge that much remains unsettled and requires further study from many disciplinary perspectives.

Finally, it is worth noting a way in which I am departing from the account of common-sense psychology that can be gleaned from Bermúdez and others, which also may constitute a (perhaps inevitable) limitation of my approach. We will recall that Bermúdez contends that *a priori* conceptual analysis cannot "tell us everything we need to know" about the kinds of concepts that are employed in both philosophical and psychological accounts of cognition (Bermúdez 2005, 10). Earlier, I said that in addition what was required was observation. Bermúdez and others are more specific, claiming that full understanding of such concepts requires experimental and other scientific investigations into the meaning of folk psychological concepts (Malle 2006, 209, Bermúdez 2005, 10-11, Guglielmo, Monroe and Malle 2009, 459), and their relationships to each other (Rips and Conrad 1989, 189).

Given what I have said about the state of psychological science in the previous section, it seems to me that an attempt to develop an account of autonomy that draws on scientific evidence would likely fall prey to many of the same problems facing situationist arguments, particularly those relating to the quality of the individual studies as well as those relating to the availability of the relevant meta-analyses. It may be, as Paul Churchland has claimed, that we find ourselves still too early in the maturation of psychology and affiliated sciences to reliably supplant common-sense psychology despite its admitted flaws as a tool for understanding social life (Bermúdez

2005, 174). Even if this is the case, and if common-sense psychology will be some day replaced by a more rigorous and advanced theoretical framework, it seems that if we proceed carefully and conservatively, our own understanding of human capabilities and our assessments of practicability and plausibility in the context of autonomy will be justified at least by the fact that they extend from the framework we use relatively reliably to understand human behaviour on a daily basis.

In the rest of this dissertation I intend to apply a common-sense standard for assessing practicability in the service of establishing the inadequacy of other positions regarding the nature of autonomy, and the adequacy of my own. Some might see this as a reliance on intuition, but my hope is that this final section of the present chapter has gone some way in showing how our sense of what people are capable of in this area of discourse is supported by more than gut feeling, it is continuous with the resources we use to successfully navigate the world. And so, while not perfect, our pre-scientific judgements about mental life have a certain amount of weight that may not be enjoyed by mere intuitions. In addition, I plan to proceed conservatively by relying on as uncontroversial claims about human mental abilities and their limits as possible. In cases where this is not easily achieved, I will draw connections to easily recognizable behaviour by way of analogy. A reliance on common-sense understandings of human psychology and abilities is not unusual among autonomy theorists, though I have yet to see others articulate what they take common-sense to involve as explicitly as I have here. It is my contention that many important developments and innovations in how we understand autonomy are the result of employing common-sense critiques of earlier conceptions.

Through the process of responding to two potential objections to my project— one that would reject taking practical considerations into account when theorizing about autonomy, and one that

rejects the plausibility of autonomy on empirical grounds— I have gone some way to articulating what kind of constraints I take to apply to theorizing about autonomy. In particular, I have defended the idea that a credible understanding of autonomy will be shaped both by our folk understanding of psychology and by the nature of the social practices in which autonomy is exercised. And, I have described the nature of these constraints as I see them. In the next chapter I will survey what I take to be instances in which a common-sense understanding of psychology has been used to identify and overcome flaws in some conceptions of autonomy. The resulting picture, in my view, has done much to enrich our understanding of the skills and capacities involved in autonomy. I will also contend that a new potential problem for autonomy emerges through this process. By the end of the next chapter I will show how this theoretical enrichment has had the unfortunate side-effect of exacerbating autonomy's problem with *extensiveness*, the number of abilities and therefore ideas it involves, which has implications for both its coherence and practicability.

Chapter 2: Common-Sense Psychology Generating Insights and Renewed Problems?

In the previous chapter, I defended the position that an adequate understanding of autonomy should be informed by empirical considerations. My argument rested on two factors: the involvement of many psychological concepts in the broader concept of autonomy, and the relationship between autonomy and social practices. I also proposed that in order for a conception of autonomy to be deemed adequate it is important for it to be not just possible to achieve the requisite level of functioning, but that the plausibility of a conception of autonomy rests at least in part on whether achieving requisite levels of functioning is practicable for people generally. Otherwise, there is considerable room to doubt that such social practices and expectations are suited to creatures with our abilities and limitations. I therefore take both psychological considerations and norms about what it means to participate in, possibly idealized versions of, social practices as constraints on how I will develop my view.

With the importance of empirical considerations established, I turned my attention to an important contemporary rejection of at least the practicability (if not possibility) of autonomy predicated on experimental findings from psychology and related fields. In explaining why this stance was at best premature, I highlighted limitations related to obtaining sufficient evidence of the right kind to meet the evidential standards that have recently emerged in these disciplines. In so doing, I revealed this to be a systemic problem that also has implications for how I am to make judgements about the practicability of autonomy in the present work. Scientific evidence supporting my view would be expected to be similarly flawed. In consequence, I adopted a suitably modified version of common-sense psychology which helps locate our pre-scientific judgements about human capacities within a framework of theoretical ideas that we all use with considerable success to understand and navigate the social world. Although this framework is

admittedly here imperfect, I have defended our reliance on it as more robust than sheer appeals to intuition.

In this chapter I will discuss developments in contemporary theorizing about autonomy in order achieve two tasks. The first task is to examine the way in which common-sense judgements about practicability are already at work in contemporary theorizing about autonomy. I noted that it appears that other autonomy theorists have long been using some sort of common-sense understanding of human psychology in their work. I would contend that concerns of a common-sense psychological nature have motivated important developments in refining our understanding of the role of capacities such as reason and introspection in autonomy, as well as in elucidating the relationship between autonomy and other features of persons such as independence. Some thinkers who have taken up the question of personal autonomy in contemporary discourse have worried that autonomy may be implausibly demanding due to the level of achievement in rationality and introspection that seems necessary to function adequately. Still others have charged that autonomy requires an implausibly high level of self-sufficiency. I will review these concerns and discuss how they have been mitigated. Thus, the first section of this chapter will examine responses to some important common-sense problems associated with autonomy and outline the developments that resulted from engaging with these issues.

My second task in this chapter is to articulate how theoretical contributions aiming to improve the alignment between our understanding of the capacities and other factors involved in the exercise of autonomy and our common-sense understanding of psychology raise renewed concerns about the viability of autonomy as a concept. I will begin to specify how, in my view, despite the productive and insightful nature of the many additions theorists have made to the autonomy picture, the added *extensiveness* of the concept generates twin problems for theories of

autonomy. The first problem associated with increased *extensiveness* is that it exacerbates existing concerns about the coherence of autonomy as a concept. The nature of this problem will be addressed briefly in this chapter, and given a full treatment in Chapter 3. The second problem with increased *extensiveness* is that it raises renewed doubts about the general practicability of the exercise of autonomy. A preliminary articulation of this problem will appear in this chapter in anticipation of an expanded treatment in Chapters 4 and 5.

Rescuing the Practicability of Autonomy Using Common-Sense Psychology

The purpose of this section is to demonstrate how common-sense psychology has been used to identify and develop responses to problems relating to the practicability of autonomy. I will address two relatively well-developed sources of concern about autonomy from this perspective. First, there are concerns about the mental demands of autonomy. Some thinkers who have taken up the question of personal autonomy in contemporary discourse have worried that autonomy may be implausibly demanding in at least two ways. In the first two subsections, I will discuss the potential excessive demands of autonomy relating to rationality and introspection, and discuss how these concerns have been addressed. Second, others have worried that autonomy requires an implausibly high level of self-sufficiency. One purpose of the discussion in this section is to show common-sense psychology at work in identifying and providing resolutions to conceptual issues related to practicability that have arisen in the process of theorizing about autonomy. The other purpose of this section is make the reader aware of important conceptual developments that should be kept in mind as our exploration of the nature of autonomy progresses.

i.   Reason and Autonomy

The first potential problem for the practicability of autonomy to be discussed in this section has

to do with the role of reason in autonomy. Reason, or reasonableness, plays an important role in

accounts of what gives agents' endorsements of self-defining attitudes weight as counting toward

their authentic perspective and therefore to their autonomy. However, theories about personal

autonomy can employ widely diverging conceptions of reason with significant differences in

consequences for the way we understand what autonomy entails. An important distinction in this

area is between procedural and substantive accounts of reason.[27] Proceduralist conceptions of

reason in the contemporary autonomy tradition might most easily be traced to Gerald Dworkin's

articulation of what kinds of endorsement are consistent with autonomy. Reason, for Dworkin,

is part of the authenticity requirement that he proposed as one aspect of autonomy. Dworkin's

view requires that people have the ability to rationally and critically evaluate their situation

(Dworkin 1976, 25) in order for their endorsements to be authentic. It seems to me this idea has

been taken up in more detail in subsequent theorizing about autonomy. For instance, Diana

Meyers claims that part of autonomy competence involves "critical rationality", which involves

identifying and assessing reasons supporting various options in order to establish their relative

merits (Meyers 1989, 81-2). And, John Christman's rationality criterion for autonomy is

characterized by an absence of inconsistencies or "mistakes in logical inference" (Christman

1991, 12). The key feature of procedural accounts of the role of reason in autonomy is the use of

---

[27] Note *both* reason *and* autonomy as I see it have procedural and substantive varieties. And, it is important to recognize that committing to one variety regarding reason does not necessarily commit you to the same variety with respect to autonomy. One might expect there is at least some correlation since accounts of autonomy that have included substantive rationality have tended (rightly or wrongly) to be labelled as at least *weakly substantive* accounts of autonomy.

acceptable forms of inference and judgement in arriving at conclusions about which self-defining attitudes or courses of action to adopt or reject.

By contrast, substantive accounts of rationality depict it as requiring the possession of a certain kind of knowledge or understanding that extends beyond just the use of sound deliberative processes. A perhaps paradigmatic example of an account of autonomy that leverages a substantive account of rationality is Susan Wolf's Sane Deep Self View. Wolf believes that overall the picture of autonomy painted by people like Harry Frankfurt and Gary Watson identifies a necessary condition for autonomy, namely the ability to govern one's actions and desires through a process of self-critical introspection and endorsement (or self-revision) (Wolf 1989, 142-43). But, she does not think this is sufficient for autonomy given the possibility that one's fundamental evaluative standpoint could be distorted, resulting in mistaken judgements about what is good and important in life (Wolf 1989, 146). Thus, she claims that autonomous people must also be *sane*; they must be able to appreciate the world for what it is, "to be controlled… by perceptions and sound reasoning that produce an accurate conception of the world rather than by blind or distorted forms of response" (Wolf 1989, 145). The standard of rationality employed by Wolf is substantive in that it is not merely the process of adequate reasoning that makes our conclusions rational, rather they must possess specific content. This content is fixed by the facts of the world, and as such the kind of reasonableness she views as important for autonomy may also be described as externalist.[28]

---

[28] The distinction between 'internalist' and 'externalist' standards of rationality would be an acceptable alternative for dividing up versions or rationality employed in the context of autonomy. I imagine that the resulting groupings of views would be somewhat different since 'externalist' and 'substantive', and 'internalist' and 'procedural' are not strictly equivalent.

On either description of rationality, concerns arise about whether people must be highly, or even perfectly, rational in order to be autonomous. From a common-sense perspective, we might have two slightly different worries about procedural and substantive rationality. In regard to the proceduralist account of rationality, it may appear as though being autonomous requires being, potentially exceedingly and reliably, smart (Christman 1991, 15). This produces two potential sub-problems. The first is that if the standard for logical scrutiny of and compatibility between our self-defining attitudes is so high as to be unrealistic for people in general, then our account would fail to reflect realities about the extent to which people generally examine their self-defining attitudes, since it would appear all but impossible, given what we all know about human fallibility, that people generally (if anyone at all) have (or has) sufficiently logically coherent desires (Christman 1991, 14). The second problem that attaches to a procedural conception of rationality is related to its weaker form in which perfect exercise of logical inference is not what is required for autonomy. Instead, we must set some appropriate threshold that establishes who counts as autonomous and who does not. The difficulty is establishing a threshold that excludes only the right people, ensuring that those who deserve moral credit and political respect are afforded it.

The problem for substantive criteria for rationality[29] from the perspective of common-sense empirical adequacy, is that, as Susan Wolf points out, autonomy potentially comes to be too

---

[29] Opponents of this approach to the rationality requirements for autonomy might argue that substantive criteria for rationality undermine the point of autonomy altogether since they demand that people believe certain things, making individuals' judgements irrelevant to advancing their autonomy (Christman 1991, 14, Berlin 2002, 180-81). This kind of objection is at the heart of what is, for me at least, a difficult theoretical point. One might set substantive criteria for autonomy in two ways. The first, which is more clearly in conflict with the underlying insight behind autonomy, is based claims that there are correct conclusions about what is valuable or right that people must arrive at in order to be autonomous. It is less clear to me that substantive rationality, in which externally defined standards for arriving at the kinds of judgments that are the appropriate basis for autonomy are articulated, are fundamentally contrary to autonomy in the same way. There may simply be certain standards that deliberation that is consistent with autonomy must meet, but this does not necessarily determine the kind of lives people must live.

closely connected "with being right about the world… implying that anyone who acts wrongly or has false beliefs about the world is therefore" not autonomous (Wolf 1989, 149). The problem that Wolf is identifying here is clear. Commonsensically, we know that people often act on the basis of false or incomplete information, and so autonomy would be extremely difficult or impossible if it relied on having perfect information. Autonomy must be consistent with the possibility of errors and false starts. And so, a practicable account of autonomy must carefully set the standard for how much needs to be known, how right we need to be about the world to reflect the capacities and circumstances of real people.

Many philosophers, while acknowledging the importance of rationality for autonomy, have been sensitive to these concerns in establishing the practicability of their own accounts. In response they have developed a number of more modest criteria to bring the rationality requirements associated with autonomy more into the reach what we take people's abilities to generally be. One of these has involved setting a requirement of only "minimal" rationality for autonomy. What this looks like depends on the conception of autonomy in use. On a proceduralist conception of rationality, minimal rationality might involve "not being guided by *manifestly inconsistent* desires or beliefs" (Christman 1991, 15), which would depend on at least some ability to make good logical inferences in considering one's means, ends, beliefs, values and so on. On a substantive conception of rationality, minimal rationality might involve "the minimally sufficient ability to cognitively and normatively recognize and appreciate the world for what it is" (Wolf 1989, 145). This version of minimal rationality requires that our perceptions are controlled by, or connected to, the environment in a way that gives us at least a roughly locally accurate picture of the world rather than one that has gaps or distortions (Wolf 1989, 144-45).

A further way in which the rationality requirement for autonomy has been made more consistent with the abilities we take normal people to have is by recognizing that rationality (on either conception at use here) is a capacity or ability. In recognizing this, it becomes possible to separate the possession of a capacity from its exercise (Wolf 1989, 150, Meyers 1989, 89). So, our conception of autonomy needs to account for the fact that autonomous people will likely at times fail to deploy their rationality in the right way at the right moment. A rationality criterion readjusted in these ways allows for fallibility and relatively modest mental capacities to be consistent with autonomy. At the same time, it still excludes at least those who are greatly and systematically deluded about the world and about the relationships between their means and ends. While these modifications to how we view the kind of rationality at play in autonomy still leave the threshold for who counts as autonomous unarticulated, I would contend that they represent important insights into what kind of rationality criterion can be reasonably incorporated into an empirically plausible conception of autonomy.

ii.   Introspection and Autonomy

As noted above, a second concern about the mental demands of autonomy relates to the extent to which autonomous people are expected to have scrutinized their actions and values. Recall that authenticity requires that people are able to choose or discover the self-defining attitudes that make up their will in a way that makes these attitudes their own "in some important sense" (Levy, 2006, p. 2).[30] Thus, the exercise of autonomy will involve arriving at one's values through processes such as approval, identification, or endorsement and directing one's life according to them. When our values or other self-defining attitudes result from an unproblematic process of

---

[30] See also Wolf 1989 p. 140.

deliberation[31], they are deemed to authentically belong to those who possess them, and therefore to be significant in the way that theorizing about autonomy seeks to specify. In response to this account of the nature of the formation of authentic self-defining attitudes, we might develop three concerns about the plausibility of autonomy from a common-sense perspective. Two of these arise from doubts about our ability to know our own motives, and the other stems from doubts about the consistency between autonomy's basis in the endorsement of our self-defining attitudes and a recognition of the limitations of human cognitive capacities.

More specifically, the first concern is that the explanation of how certain desires and behaviours come to be autonomously held suggests that our motivations and the ways in which we come to acquire them are transparent to us (Christman 1991, 17). Autonomy only seems possible if we can be consciously aware of, and accurately identify, the self-defining attitudes that motivate us. This is one of the practical implications of autonomy that situationist arguments, and the research that they build on, bring into question. As we have seen, the relevance of situationist arguments to contemporary theories of autonomy is that they raise questions about the extent to which people are able (1) to be aware of the forces that actually motivate them to act and (2) to identify the causes of their actions correctly. At best, doubts on these points makes our capacity for autonomy seem more limited than we would expect. At worst, they suggest that autonomy is not actually possible.

In Chapter 1, I argued that the empirical basis for situationist arguments is weak and so the view does not represent a credible challenge to autonomy as presented. Further resistance to the stronger versions of situationism from the perspective of common-sense psychology could also

---

[31] Briefly this process at minimum must be sufficiently rational in the ways discussed above and not subject to the manipulative or coercive actions and influences of others. There are likely other criteria that must be met as well to be discussed in later sections.

simply point to our common-sense practice of attributing knowledge of our reasons for acting to ourselves and others. This is not supposed to suggest that people have perfect knowledge of why they choose and act they way they do; a common-sensical view of human psychology can acknowledge that our capacities for introspection might be limited or imperfect in certain ways. For example, we might be more accurate the more time we have to deliberate and reflect on our reasons.[32] As such, a common-sense psychological approach to evaluating practicability, although it would rule out proposals that discredit the idea of autonomy altogether, must take the prospect of limitations in people's capacities for introspection seriously.

Relatedly, the second concern one might have about the practicability of autonomy insofar as it requires us to know ourselves is whether to be autonomous someone must have introspected so thoroughly as to have approved of all of their self-defining attitudes. This might seem to be the case given that it is theorized that people's attitudes towards their desires are what make these self-defining attitudes *theirs* (Dworkin 1976, 25). This suggests that all authentically held self-defining attitudes must have been consciously approved. Presumably, people's autonomy is undermined "to the extent" that their self-defining attitudes remain unexamined (Meyers 1989, 28). Yet, a common-sense psychological outlook might note that "few of us have examined *all* our beliefs and preferences[33]" (Christman 1991, 14), and doubt that we ever could. At bottom, we may worry that exercising autonomy involves so much introspection as to lead to a hyper-intellectualized, or mentalistic existence (Meyers 1989, 83, Meyers 2005, 27). But, if an account of autonomy were to contend that exercising autonomy requires such a high degree of introspection and self-knowledge, from a common-sense psychological perspective it would be

---

[32] Though I will propose that there may be some ways to avoid some of the autonomy-related problems associated with time constraints in Chapter 4.

[33] Emphasis mine.

judged to be so stringent and time consuming as to be implausible (Christman 1991, 14, Meyers 1989, 84). To contend with this, existing theories of autonomy have done several things to demonstrate that autonomy is not as cognitively demanding as critics allege on the point of self-knowledge.

This first approach to defending autonomy against the charge of being excessively demanding in this sense is to defend the view that autonomy is consistent with error and with imperfect self-knowledge. A good example of this first strategy appears in Diana Meyers's *Self, Society, and Personal Choice.* On her view, the introspective processes that confer autonomy are never finished, so she accepts that people's knowledge of their values and goals may be imperfect and that desires that were at one point endorsed may at some point be rejected or revised (Meyers 1989, 49). But, she also views the process of coming to know one's values and goals as an active and experimental process in which introspection may not be enough (Meyers 1989, 83). Trial and error will therefore also be needed so that people have a chance to try out different ways of life and determine what is right for them on a more concrete basis. On her view, authenticity "is a moving target" (Meyers 1989, 90) and the experimental process is part of the exercise of autonomy (Meyers 1989, 83). On Meyers's picture, we cannot and should not expect people to have comprehensively engaged in a definitive process of introspection. Rather, this is an on-going and organic process.

The second approach is to deny that the process of self-knowledge is entirely conscious or rational. Here I will highlight three ways in which this idea has been developed. First, we may come to know and define ourselves through emotional processes. Thus, the burden of self-knowledge and self-definition does not fall solely on our reasoning capacities. While the ways that we arrive at some of our values may involve deliberation, sometimes these processes will be

guided by affective responses instead (Killmister 2013, 110). For example, the detection of conflicts or other problems among one's self-defining attitudes might be achieved through observing when one experiences disturbing feelings (Meyers 1989, 54) such as chronic regret (Meyers 1989, 35) "[s]hame, dismay, exasperation, or disgust with oneself" (Meyers 1989, 86). According to Meyers, such negative feelings can signal failures in exercising autonomy, which upon investigation may reveal unknown values or internal conflicts. Similarly, the process of adopting new goals, values, or undertakings maybe as much a process of trying them and seeing how they *feel* as reasoning through them.

The second way in which processes of self-definition and self-knowledge may not be rational or conscious is by being subconscious (Killmister 2013, 110). Autonomous people may be able to act spontaneously in accordance with their authentic values without consistently reconsidering them (Meyers 1989, 86). Proponents of this view would argue that autonomy is largely "a background phenomenon" (Meyers 1989, 55).

A third way in which to defend the position that the processes of self-definition and self-knowledge are not liable to exceed our mental powers is to argue that these processes can be enhanced by our relationships with others (Mackenzie and Stoljar 2000, 22). In particular, we might come to know ourselves better through our relationships with others. When we are confused at our emotional responses to, or evaluations of, candidate self-defining attitudes, an outside perspective might provide a new interpretation of our situation and the reasons for our responses (Meyers 1989, 79). So, while we might not always have the cognitive resources to know ourselves, we can improve our capacities for self-knowledge by leveraging the perspective and cognitive resources of those around us.

The third and final method for defending the broader practicability of autonomy in the face of the reality that people often do not know all their self-defining attitudes is to argue that the truth of certain counterfactual claims negates the need for further (self-) knowledge. Two examples of this strategy appear in a qualification made to John Christman's historical account of autonomy, and in Wayne Sumner's subjective metric for how much information is required to take seriously an agent's satisfaction with her life. According to Christman, an important consideration in determining whether a self-defining attitude is autonomous is the process through which it was formed (Christman 1991, 10). There are hints of this idea at least as far back as Dworkin's "Autonomy and Behavior Control", in which Dworkin discusses how our desires may or may not be autonomous based on our attitudes towards their sources. For example, we might either reaffirm or repudiate our desire to ski after discovering that it is inspired by our envy towards an athletic sibling (Dworkin 1976, 25). So, it is not just a desire itself that needs endorsement for the purpose of establishing its authenticity. On Christman's model, it is primarily the process or conditions of desire formation (rather than underlying attitudes) that is of interest from the perspective of autonomy (Christman 1991, 10).

In shifting the focus of assessments of autonomy to these considerations, two central questions become "whether [an agent] resists [the] process when (or if) given the chance" and whether the process by which certain desires were formed even allow the possibility of judging the acceptability of the formative processes (Christman 1991, 10). One of the criteria that Christman offers for assessing autonomy in this way requires that the agent "would not have resisted that process, had" she attended to it (Christman 1991, 11). While self-awareness is an important faculty supporting autonomy on Christman's view, by stating his historical criterion for autonomy as a hypothetical he allows for a counterfactual test for authenticity. According to this

test, even if an agent remains unaware of the process by which their desire has been formed, if we have reason to believe that they would have accepted the process then having the desire is at least consistent with them being autonomous.

The example derived from Sumner is not as closely related to autonomy as the above historical criterion from Christman. It relates primarily to the relationship between our appraisals of our own happiness and welfare. But, it is relevant to autonomy because it concerns the appraisals of our lives in a way similar to what is required for authenticity. By Sumner's lights, requiring that people's assessment of the conditions of their lives be true or even justified in order for their assessment of their own happiness to be deemed a reflection of their well-being is unrealistic and arrogant (Sumner 1996, 158-59). This would make it possible only for rational (and well informed) people to have authentic well-being. But since our knowledge of ourselves and the world[34] is necessarily limited, Sumner suggests that having further information is only important if it would make a difference to how we view our lives (Sumner 1996, 160). Thus, Sumner argues that people's assessments of whether they are satisfied are to be taken as authoritative, even if these assessments are incomplete or in error, unless we have reason to believe that the process by which they have made this assessment is not procedurally independent (Sumner 1996, 171), or we believe that further introspection would result in a re-evaluation of how happy or satisfied they were with life (Sumner 1996, 161).

Sumner's claim about happiness can also be applied to autonomy. While there are important differences between welfare and autonomy, the general insight-- that further knowledge is not necessary if it will not affect the agent's judgment-- would appear to be plausible for both cases.

---

[34] Sumner's account has mostly to do with our knowledge of the world, I am extending his argument to cover self-knowledge.

Thus, both Christman and Sumner's arguments provide grounds for believing that a person can be autonomous even if their introspection is incomplete unless some countervailing condition is thought to prevail (i.e. that a person would have resisted the process by which her desires were formed, or that she would change her desires based on more complete introspection).

A final concern about the mental demands of autonomy as they relate to introspection is that some might believe that exercising autonomy adequately requires that we constantly consciously evaluate and choose our actions. This would seem to both conflict with the reality that often our actions are unreflective, as well as deny the consistency between spontaneous or habitual action and autonomy. Indeed, as Meyers remarks, an overly strong commitment to deliberate decision-making as fundamental to the exercise of autonomy threatens to paint autonomous people as "inhibited, rigid, unspontaneous, and shallow—in a word, inhuman" (Meyers 2005, 29). And this, from the perspective of a common-sense approach to psychology, would be deeply problematic since it not only fails to match our general knowledge or understanding of mental life, but also paints quite an alien picture of what it is to be autonomous.

Common-sense psychology would therefore privilege accounts of autonomy where it is not undermined by the fact of automatic behaviour. A number of accounts of autonomy have taken up this challenge. For instance, Frankfurt himself acknowledged that the alignment between someone's will and their higher-order volitions can be spontaneous and at times achieved without thought (Frankfurt 1971, 17). He notes that while some people may have to struggle to live in accordance with the values and desires with which they identify, for many others acting out one's volitions is entirely natural and easy. As such, we don't have to believe that the mental lives of autonomous people are consumed with the exercise of continually re-evaluating their motives and carefully selecting their actions (Meyers 1989, 86). This suggests that autonomy is

not only consistent with automatic actions, but that these actions can actually be expressions of one's autonomy. This point becomes clearer when we recognize that introspection is a capacity, and like all capacities it may be something we possess, but that we do no exercise all the time (Meyers 1989, 86). The question then becomes how spontaneous action, which does not rely on the exercise of our introspective and rational capacities, could be an expression of our autonomy.

At least three mechanisms can be identified as establishing the continuity between autonomy and spontaneity: natural inclination, habit, and the operation of unconscious "agentic powers" (Meyers 1989, 86, Meyers 2005, 33). Natural inclination presumably refers to people's endorsement of self-defining attitudes that they have ended up with through biological factors or through socialization, which is more likely the source of these commitments. In such cases, action stemming from these inclinations will also be autonomous though potentially unreflective. Through the cultivation of habits that derive from our values and goals, we might make certain responses or actions into inclinations that operate much the same way as natural ones. Finally, on a more comprehensive understanding of the skills required for autonomy (which we will be developing in a later chapter), it may come to light that some of these operate at least partly unconsciously, or occur outside consciousness (Meyers 2005, 39, 33). The possibility of sub-conscious or habitual exercise of autonomy provides a way of avoiding worries about the extent to which autonomous agents should be expected to be robotic and rehearsed. It also provides a way of understanding how in many cases autonomy is "a background phenomenon" (Meyers 1989, 55). On Meyers's view autonomy involves not constant introspection and deliberation, but rather ongoing alertness and attentiveness to affective cues (awkwardness, guilt, disgust) that signal we have strayed from our authentically held self-defining attitudes (Meyers 1989, 91).

The way in which introspection is required for an adequate exercise of autonomy features in a number of objections raised against autonomy from the perspective of a common-sense understanding of human capacities. But, as we have seen, this common-sense framework can also be appealed to in resisting these very objections. Arguments in this vein have provided convincing reasons to believe that autonomy is compatible with imperfect self-knowledge and with the limits of our introspective abilities. They have also suggested aspects of mental life that are recognizable by a common-sense understanding of human psychology that help establish the consistency between autonomy and automatic behaviour. In the final sub-section of this section, I will show how common-sense concerns about the self-sufficiency demands of autonomy have been contended with from the perspective of common-sense ideas about practicability.

iii.   *Self-Sufficiency Demands of Autonomy*

A final recurring common-sense concern about autonomy that I wish to address is about the relationship between autonomy and independence. The worry is well developed in Isaiah Berlin's critique of positive freedom. The homology between the way Berlin describes positive freedom and contemporary conceptions of autonomy is apparent in his characterization of positive freedom as the wish "to be moved by reasons, by conscious purposes, which are my own… of conceiving goals and purposes of my own and realising them" (Berlin 2002, 179). This is much the same as the kind of activity envisioned by conceptions of autonomy that involve living in accordance with one's authentic values. As we saw earlier, Berlin takes the implication of such self-mastery language to be that positive freedom involves neither being a slave of people, passions, or nature (Berlin 2002, 180). In addition to having metaphysical concerns, Berlin is skeptical about how freedom from other men is to be achieved.

On Berlin's account, one of the dangers of pursuing positive freedom is that it may lead people to believe that what is necessary in order to avoid "being crushed" by the laws, activities, and accidents of both men and nature is a "strategic retreat into an inner citadel", which involves withdrawing into the mind and abandoning society (Berlin 2002, 183). Berlin claims that the result of pursuing positive freedom as an ideal of autonomy is reflected in the individualistic image of a solitary sage, who has freed himself from the influences of others by retreating into his mental world (Berlin 2002, 186). While such a person "care[s] no longer for any of [society's] values… [is] isolated and independent, on its edges, no longer vulnerable to its weapons", Berlin worries that such a life smacks heavily of self-defeating isolationism (Berlin 2002, 182). Similar conceptions of autonomy have been the target of criticism from feminist philosophers, who reject the "substantive independence and self-sufficiency" at least traditionally associated with autonomy (Stoljar 2000, 94). Lorraine Code, for example, charges that valuing and pursuing autonomy perpetuates individualistic ideals like self-sufficiency, self-reliance, and social transcendence that are not only unrealistic, but also have troubling epistemological, moral, and political implications (Code 2000, 183). Though Berlin also has concerns about the political implications of the image of the self-mastering person (Berlin 2002, 182), I will focus on the ways in which Berlin and feminist scholars believe that the kind of self-sufficiency or independence implicated in autonomy is unrealistic.

Berlin believes that in attempting to free themselves from the power of social and natural forces people adopt an understanding of necessity through which their desires and aspirations are constrained to follow only the path of reason (Berlin 2002, 189). In much the same vein as the concerns about rationality above, Berlin is concerned that this vision of autonomy demands a kind of perfection (Berlin 2002, 200) that is likely to be at odds with (Berlin 2002, 182), or

unachievable by (Berlin 2002, 197), our empirical selves. But, I think he would also be sympathetic to feminist critiques of the individualistic interpretation of autonomy, which takes other people and society to constitute a threat from which people must protect themselves (Mackenzie and Stoljar 2000, 6). This idea of what it means to be a person denies the importance of social cooperation and community. It also fails to recognize the role that other people play in establishing and sustaining our identities and personal abilities (Mackenzie and Stoljar 2000, 6, Berlin 2002, 202).

The idea that other people play an important role in the exercise of autonomy, that it is relational, has been developed in at least two principal senses. The first sense stresses how our relationships to others and our society more generally are the sources of our beliefs and desires. That is, our personalities are the causal product of the society and historical moment in which we are raised (Mackenzie and Stoljar 2000, 22). The second sense claims that people are socially constituted; parts of our social world participate in our self-perceptions and our identities. As part of the self, autonomy is therefore also believed to consist, at least partly, of "elements of [our] social context" (Mackenzie and Stoljar 2000, 22). In both senses, an appreciation of the role of social influences on or in autonomy undermines the plausibility of the individualistic, self-reliant, radically self-sufficient autonomous agent.

Although the highly independent figure portrayed above has been the target of considerable critique, at least in some quarters it is recognized as a caricature or exaggeration, albeit one taken to have considerable cultural power (Mackenzie and Stoljar 2000, 5, Code 2000, 183). In *Self, Society and Personal Choice,* Meyers claimed that the accounts of autonomy advanced by many of her contemporaries such as Young, Benn, Frankfurt, Dworkin, and Watson were still attempting to demonstrate how "people can elude socialization" or "transcend the impact of

69

social causes" (Meyers 1989, 40). But, the conception of the autonomous agent in use in these works hardly resembles the implausible character that has been the object of so much criticism. The autonomous agent that forms the basis for personal autonomy in contemporary discourse is fallible (Frankfurt 1988, 169). And, there is wide recognition of the fact that our values are inculcated in us through socialization (Dworkin 1976, 24, Wolf 1989, 143). On this view, it is not important that someone be the origin of his own values and desires (Frankfurt 1988, 171). Rather, it is his attitude toward these self-defining attitudes (values, desires, etc.) and potentially his attitude toward their origins that is important. This is consistent with the reality of socialization. Indeed, the idea of an isolated self that chooses in a way that is entirely separate from the ways it was shaped by others has long been ridiculed (Dworkin 1976, 24, Wolf 1989, 144). Admittedly, the account that develops out of papers like "Autonomy and Behaviour Control" and "Identification and Wholeheartedness" is incomplete. However, from our common-sense concern with plausibility, it is my view that we need not become exercised over the individualistic conception of autonomy since we will be working within a tradition that has been interested in reconciling autonomy and sociality for quite some time.

In an effort to defend the practicability of autonomy and to demonstrate the pervasiveness of common-sense psychological approaches in contemporary theorizing about autonomy, this section has considered several common-sense worries about the achievability of autonomy based on what it might demand of people in terms of rationality, introspection, and independence. In response to concerns raised on each of these points, I have shown how contemporary autonomy theorists have contended with these issues. This process has helped further clarify and elucidate the notion of autonomy that we will be using moving forward. That is, we have seen that there is good reason to believe that to exercise autonomy we do not need to have perfect rational or

introspective capacities, and that exercising autonomy does not demand independence from other people. These insights help to reconcile autonomy with human limitations and realities, making it more consistent with our common-sense understanding of human psychology and therefore more realistic and plausible. In the next, and final, section of this chapter I will proceed to articulate what I take to be an emerging threat to the plausibility of autonomy stemming from well-intentioned efforts to articulate an even more realistic conception of autonomy. This work has aimed to give a more complete enumeration of the psychological capacities that are employed in the exercise of autonomy, moving beyond conceptions that take reason to be the primary capacity among only a small number of capacities implicated (Mackenzie and Stoljar 2000, 21).

A New Common-Sense Concern

The objections addressed in the previous section relating to what autonomy requires in terms of rational and introspective capacities, as well as independence from others are central questions for what might be call the basic proceduralist view of autonomy. On this conception the central process in the exercise of autonomy is the endorsement of one's values and desires, and thus rationality[35] and self-awareness emerge as the key capacities required for autonomy. Both are important to the process of selecting or discovering values, desires, and goals in a way that reflects one's own perspective free of manipulation or distortion by alien forces (Dworkin 1976, 24, Christman 1991, 13 & 16). Meeting certain thresholds of rationality and self-awareness are what make self-reflection adequate for the purpose of establishing the special kind of relationship between agents and their desires and values that is of interest for autonomy (Christman 1991, 13). Some proponents of the proceduralist model suggest that autonomy is

---

[35] Rationality appears to be part of both authenticity and competency conditions.

primarily, or only, a "property of preference or desire formation" (Christman 1989, 13), and those that are deemed to be non-autonomous are those who fail to possess rationality and self-awareness. As noted in Chapter 1, the kinds of people excluded from qualifying as autonomous on this account include people with severe mental disabilities and low cognitive function, infants, people with severe mental illness, people in comas (Feinberg 1989, 28), as well as, people with unwanted addictions and people exposed to brainwashing or other forms of manipulation (Christman 2014, 374).

While many contemporary theories of autonomy acknowledge the importance of rationality and self-awareness as *authenticity* conditions for autonomy, at least two problems with the basic proceduralist approach to autonomy have been taken up by feminist and psychologically-oriented philosophers. First, there is reason to deny that rationality and self-awareness exhaust the set of skills that are required for authentic endorsement of one's values. Second, the *exercise* of autonomy has come to be understood as more complicated than a consideration of only the authenticity conditions for autonomy would suggest (Christman 2014, 374). That is, if we take autonomy to involve identifying one's values *and* living according to them, then autonomy requires authenticity *in addition to* skills and other bases for acting on our values.

Although different theories of autonomy subdivide the personal characteristics required for autonomy differently, the basic procedural picture has been amended in at least three ways[36] in light of the considerations above. First, more capacities have been identified as underlying authenticity. Second, capacities underlying our ability to act on our values have been proposed. And third, a set of attitudes that contribute both to authenticity and competence have been

---

[36] See Meyers 1989 versus Mackenzie 2014, for example.

advanced. These contributions to the model of autonomy that traces back to Dworkin and Frankfurt not only increase its nuance and richness, but also move us further away from a purely procedural account of autonomy. In what follows, I will outline some of the ways in which the basic proceduralist model has been elaborated upon in each of these three areas. The aim here is not to conduct a comprehensive survey, but rather to provide a sense of the kinds of additions proposed and the reasons for them. Once we have a sense of the ways in which the basic procedural picture has been enriched, we will be prepared to see how these changes raise renewed questions about the practicability of autonomy.

i. *Capacities for Authenticity: Beyond Rationality and Self-Awareness*

Some of the additional capacities that have been proposed as important for establishing authentic self-defining attitudes include: reasonably accurate perception, the ability to take critical distance, the ability to read or be attuned to emotion, and imagination. Arguments defending or justifying the importance of each of these to authenticity will be presented below.

As discussed earlier in this chapter, Wolf was an early critic of the idea that reflective endorsement of goals, desires, values, etc. (which presumably includes authentic, uncoerced affirmation thereof) is enough to establish that they belong to those who endorse them in the way that is important for autonomy (Wolf 1989, 140). Recall that her worry was that our deepest selves —on which our valuational and evaluative processes rest— are not fully characterized by a rationality and reflectiveness free of coercion (Wolf 1989, 143). She contended that these powers and forms of control were not enough to guarantee authenticity because they were not enough to ensure the autonomy of people who had been exposed to kinds of socialization that distort how the world is perceived and understood (Wolf 1989, 145). This led her to advance the substantive conception of rationality discussed earlier—we must be under the control of the

world (on her view) in such a way as to be able to "appreciate the world for what it is" (Wolf 1989, 145). While Wolf's substantive account of rationality remains controversial, the idea that autonomy relies on an upbringing that is not deeply distorting of our way of understanding the world, and that allows people to take a critical distance from our upbringing is similarly recognized in the inclusion of historical factors as relevant to our judgements about autonomy (Christman 2014, 374). Thus, reasonably accurate perceptual skills and the ability to take critical distance from our values are important further candidate capacities for authenticity.

A host of further "autonomy competency"[37] skills have been proposed by Meyers as the bases of the three functions she believes comprise autonomy: "self-discovery, self-definition, and self-direction" (Meyers 1989, 76). In addition to rationality and introspection, endorsing one's values authentically as part of self-definition, according to Meyers, involves skills like the ability to read feelings, and imagination (Meyers 1989, 78-9, Meyers 2000, 166).[38] The ability to read feelings is perhaps best described as a component of introspection that is underrecognized on accounts of autonomy that are highly rationalistic. On a more rationalistic account, self-awareness might mean merely being aware of one's "motivating desires and beliefs" and not being the subject of self-deception, dissociation, or identity fragmentation. (Christman 1991, 17). By contrast, Meyers has argued that decisions about the values or goals we should endorse likely exceed the boundaries of formal or practical reasoning (Meyers 1989, 79). As such, people rely by necessity on affective cues like frustration, attraction, and aversion as we consider which ways of life we should pursue (Meyers 1989, 79, 81). Our feelings about real or imagined options, or about our experiments in living, are important guides, on her view, as to whether

---

[37] I don't think the authenticity/competency divide as it currently exists is present in *Self, Society, and Personal Choice*. Meyers's use of 'competency' refers to both as far as I can tell.
[38] There are others, but I will focus on these two here.

certain goals and values are desirable and as to whether we are living authentically. And, thus she claims that an attunement to these emotions is an important capacity supporting authenticity.

The foregoing also gestures toward the important role imagination plays in selecting which courses in life we might want to follow. Without the ability mentally to represent the options to ourselves in a reasonably vivid and accurate way, we will not be able to assess them in a way that is sufficiently robust to justify endorsement (Foddy & Savulescu, 2010b, p. 35). Thus, the authority of our endorsements relies in part on the quality of our imaginations. In addition to this, imagination is important in coming to articulate to ourselves what values, goals, and desires might be options for us to begin with. The ability to draw on a "rich and complex" imaginative repertoire, and giving one's self permission to do so, is also extremely important for authenticity in self-determination (Mackenzie 2000, 129). If our imaginings are restricted by placing too much importance on things that are not essential, or by failing to consider something that, in retrospect, should have been obvious, the boundaries of our imaginings become restricted. This restriction threatens to produce an unreliable result in what we deem to be valuable, possible, or consistent with our self-understanding. Ensuring that imagination is not overly constrained helps avoid self-limiting and other negative outcomes, making it more likely that we will choose authentically.

ii.   *Competency Skills*

As noted above, not only has the set of skills implicated in authenticity been expanded, but additional competency skills involved in acting out our authentic values and desires have been put forward. Prime among these are volitional (Meyers 1989, 83, Mackenzie 2014, 18) and motivational capacities (Christman 2014, 373). With respect to the volitional capacities, self-control is required to exercise autonomy, not only in opting to act in accordance with values, but

also in being able to avoid distraction or self-sabotage during the pursuit of goals and desires (Dubljevic 2013, 48, Levy 2006, 9-10). So, self-control involves keeping oneself on track in the face of temptations or demoralization. Other volitional capacities may include competencies Meyers has articulated under the terms 'resistance' and 'resolve'. These relate to the kinds of social opposition that people might face in living according to their authentic values. Demonstrating resolve involves acting even in defiance of social norms or the scrutiny of others, resistance involves persisting in the face of unjustified pressure or opposition from others (Meyers 1989, 83-4).

Motivational competencies, as I read them, are the skills that people need in order to be moved by their values and desires to act. These might include capacities like the ability to appreciate the force of good reasons, motivational effectiveness, and decisiveness (Mackenzie 2008, 520, Mackenzie 2014, 33-4). With respect to the former, in order to live according to their values, autonomous people will need to be able to change their behaviour and plans in response to new reasons (Foddy & Savulescu, 2010a, p. 12). Although the change might be executed through the use of self-control, this will presumably be preceded by change in motivation resulting from a re-assessment of the situation. Motivational effectiveness involves being able to experience motivation in a consistent and reliable way (Mackenzie 2008, 520). This would exclude being weak-willed or impulsive. We might see the importance of decisiveness at work in Frankfurt's account of identification. It is possible to continue deliberating about options indefinitely (Frankfurt 1988, 167). Although the ideal may be to make decisions when we are confident our result is correct, decisions made based on reasons having to do with the cost of further inquiry can be made in the unreserved sense Frankfurt employs (Frankfurt 1988, 168). A major cost of continuing to deliberate endlessly is that it results in inaction and directionlessness. Thus, the

ability to at least appreciate when further deliberation is unreasonable, and a disposition to be

motivated by the most reasonable alternative helps avoid a pathological form of stagnancy which

would be inconsistent with autonomy.

iii.   *Attitudes for Autonomy*

The capacities listed in the previous sub-sections largely represent elaborations on, rather than

departures from the basic proceduralist model.[39] That is, so far the discussion has been restricted

to the mechanisms that need to be recruited in order for authentic endorsements to be made and

carried out. However, many philosophers, especially those in the feminist tradition, have argued

convincingly that autonomy cannot be exercised, nor important capacities that participate in it

deployed, without the presence of several self-regarding attitudes in would-be autonomous

agents. This is a significant departure from proceduralist thinking because it goes beyond

enumerating skills to specifying what people need to believe during the deliberative process in

order for it to be effective. As a rule, proceduralist theories claim to be "content neutral" (Levy

2006, 1, Mackenzie and Stoljar 2000, 13, Mackenzie 2008, 519). Thus, theories of autonomy that

include attitudes qualify either as not purely procedural, or weakly substantive (Stoljar 2000, 95,

Mackenzie 2008, 512). The attitudes that have been advanced are claimed either to constitute

pre-conditions for the use of certain autonomy skills (Mackenzie 2008, 525) or to defend against

the deleterious effects of socialization with which purely procedural accounts cannot contend

(Mackenzie and Stoljar 2000, 19). Some of the proposed attitudes include: self-trust/self-

confidence (Mackenzie 2008, 525, Govier 1993, 104), self-worth/self-esteem (Stoljar 2000, 95,

Mackenzie 2008, 525, Benson 1994, 661), self-respect (Mackenzie 2008, 525).

---

[39] The main departure would be Wolf's substantive account of rationality, which is usually taken to make her account of autonomy substantive as well.

According to Trudy Govier, *self-trust* involves self-acceptance, self-confidence, a certain optimism about oneself, and "a sense of one's own competence especially as regards to control, judgment, and adaptability" (Govier 1993, 101). An important aspect of this attitude is that it involves the belief in one's own dependability. She argues that procedural autonomy, especially the version articulated by Diana Meyers, depends on self-trust. She claims that self-trust is extremely important in the processes of self-definition and self-direction.

In the process of self-definition we have to introspect to arrive at some conclusions about what is important to us. Possessing self-trust is important as we make judgments (Govier 1993, 106). These judgments might be about what standards to employ when dealing with uncertainty (see Frankfurt on decisions), or they could be judgments about what to endorse. In making such judgements we must be able to trust that we have fairly and thoroughly considered the information we have and have followed it to the most appropriate or reasonable conclusion (Mackenzie 2008, 525, McLeod and Sherwin 2000, 262). Govier contends that "if we are insecure in our sense of our own values, motives, and capacities, we cannot think … effectively" (Govier 1993, 106). Nor on her view, can we act, that is be self-directing, effectively if we are not self-trusting. We must be able to have confidence in our processes for choosing our actions based on our self-defining attitudes and in the decisions themselves (Govier 1993, 108, McLeod and Sherwin 2000, 262). And, in the most extreme case, we need to be able to withstand the criticisms or questioning of our capacities by others (Govier 1993, 106). This suggests that self-trust is an important component of resistance and resolve. As such, not only our introspective and practical judgment, but also our volitional capacities depend on the presence of particular attitudes for their effectiveness. As a general statement of the idea, McLeod and Sherwin have

proposed that the exercise of the competence conditions of autonomy requires, as part thereof or in addition to, that an agent feel she *actually is* competent (McLeod & Sherwin, 2000, p. 266).

Self-respect involves thinking of oneself as an authoritative source of reasons and as having equal status and standing relative to others (Mackenzie 2008, 525). It is the sense of one's own dignity and value as a person (Govier 1993, 110). Self-respect is important in one's exercise of autonomy capacities for similar reasons to self-trust. It is important both to the process of settling on goals and projects, and to carrying these goals forward in the face of the needs and values of others. That is, it is through self-respect that the endorsements that a person makes come to have significance to her. A self-respecting person will see her decisions as having weight and her plans as worthwhile (Govier 1993, 110). It also involves an interpersonal aspect in that self-respect provides reason for people to believe that their interests, values, etc. should be given due consideration by other people. The role of self-respect in autonomy competence might be most easily discerned by considering the implications of a lack of self-respect. From the perspective of the autonomy competencies that require making judgments, a lack of self-respect is problematic because a person experiencing a deficit in this area may have trouble committing to any particular way of life. From the perspective of volitional competence, a person with little self-respect will be overly willing to give up her goals in favour of the interests of others, and will be easily dissuaded from pursuing her goals in the face of opposition or obstacles since she has little sense that her goals are *worth* the effort (Govier 1993, 110).

Finally, *self-esteem/self-worth*, while seeming very similar to self-respect is, according to Govier, a more fundamental attitude or sense. Self-esteem involves a more global positive attitude towards one's worthiness and adequacy as a person (Govier 1993, 113). It may also include a similarly generally positive stance towards "one's life, one's commitments, and one's

undertakings as meaningful, worthwhile, and valuable" (Mackenzie 2008, 525). Govier claims that self-esteem is essentially the same as self-acceptance. As such, per the characterization of self-trust provided above, the attitude of positive self-esteem is a component of self-trust and therefore contributes to autonomy through its relationship with self-trust (Govier 1993, 114).

I would argue that the development of the requirements for autonomy in the ways articulated above, concerning how authenticity is achieved, how autonomous action is carried out, and what attitudes an agent might need during these processes are, in general, salutary contributions to our understanding of autonomy. Autonomy appears to be a complex achievement, undergirded by a range of capacities that authenticate value and spur action according to them. The arguments for the need to adopt self-regarding attitudes outlined above might be more contentious. For example, Suzy Killmister has argued that these attitudes are at most correlated with autonomy, but are not necessary for it or constitutive features of it (Killmister 2013, 119). Despite Killmister's doubts, the strength of the proposal that the attitudes listed above (or something like them) are necessary for autonomy might be appreciated by considering what an agent without them would be like.

How could a person with no self-trust or respect (or who is neutral in this regard) come to establish and sustain the kinds of commitments necessary for autonomy? It seems to me that these attitudes speak to potentially hidden aspects of practical reason. If theorists about action move too quickly between an agent's thinking something is important (or wanting something), to deciding that s/he should do it, they risk concealing important assumptions about the agent. I would contend that part of what justifies this transition are often unstated claims about the agent's impressions of his or her reliability in choosing what is important (self-trust and perhaps self-worth), and his or her authority to make judgements in this arena (self-respect). Without the

help of these attitudes, and especially with the opposite attitudes, it would appear that the move

from valuing to intending or doing isn't justifiable. We might expect the consequence of lacking

these attitudes to be that our ability to exercise autonomy would be paralysed by uncertainty,

indecision, apathy, or deference to others. So, there seem to be good reasons to adopt these kinds

of additions to the picture of what is involved in exercising autonomy.

In acknowledging the value of these additions, I am brought to the point at which my renewed

concerns about the coherence and practicability of autonomy gain traction. In the final section of

this chapter I will give a preliminary account of the problems I think emerge from the increased

conceptual *extensiveness* of autonomy resulting from contemporary efforts add further realism to

our understanding of the concept.

The Problem

The previous three sub-sections represent only a sampling of the capacities, and arguments

defending them, that have been proposed as additions to our picture of autonomy. They, I think,

represent quite plausible contributions. What the previous subsections conceal, however, is the

sheer variety of additional capacities, beyond rationality and self-awareness that have been

proposed over recent years. A more complete, but certainly not exhaustive list would look

something like the following:

- Rationality (Dworkin 1976, 27, Christman 1991, 16)

- Comprehension (Mackenzie 2014, 19)

- Judgement (Dworkin 1976, 25)

- Deliberation (Meyers 1989, 55)

- (Critical) Reflection (Dworkin 1976, 26, Meyers 1989, 81)

- Self-respect (Mackenzie 2008, 525, Wolf 1989, 144-45)

- Introspection (Meyers 1989, 47, 79)

- Retrospection (Meyers 1989, 53)

- Emotional attunement to self and others (Meyers 2005, 46, Mackenzie 2014, 34)

- Sanity (Meyers 1989, 62, Wolf 1989)

- Imagination (Meyers 1989, 78, Foddy and Savulescu 2010b, 35)

- Memory (Meyers 1989, 79, Meyers 2000, 166)

- Volition (Meyers 1989, 79, Mackenzie 2014, 18)

- Resistance to external pressures (Meyers 1989, 83)

- Determination (Meyers 1989, 83)

- Curiosity (Meyers 1989, 87)

- Communicative abilities (Meyers 1989, 87, Meyers 2000, 166)

- Reliable perceptual abilities (Killmister 2013, 112)

- Interpersonal skills (Meyers 2005, 38, Mackenzie 2014, 34)

- Embodied practical intelligence (Meyers 2005, 40)

- Decisiveness (Killmister 2013, 113)

- Self-control (Killmister 2013, 114)

- Motivational competence (Christman 2014, 373)

- Self-trust (Killmister 2013, 115, Mackenzie 2008, 525)

- Self-awareness (Christman 1991, 17)

- Self-worth (Govier 1993, 110)

- Dialogical skills (Mackenzie 2014, 34)

- Disposition to answer for one's self (Westlund 2003, Benson 2005)

The citations offered in the list point to examples of authors who have proposed or elucidated each capacity, they do not represent a comprehensive catalogue of the discussions about the capacities as they relate to autonomy. Readers can consult the sources cited if they would like to see the arguments that support the inclusion of the capacities and attitudes for which a defense has not been given in this text. In my view, there are good grounds for including most of these in how we understand the exercise of autonomy.[40] And if so, then it looks like autonomy should be expected to include a wide range of abilities in both a conceptual and practical sense. It is in being confronted with this conceptual and practical *extensiveness* that I think renewed concern about the usefulness of autonomy emerges. The fact that so many different capacities are involved in autonomy raises questions about both autonomy's coherence and practicability.

In the first section of this chapter, I outlined several common-sense concerns about the plausibility of autonomy. A common preliminary concern about the rationality and introspection criteria was that perfect mastery and exercise of these capacities-- in the form of perfect inferences, perfect perception of the world, perfect self-knowledge, or perfect self-monitoring-- is clearly not achievable for people in general. Views that proposed that autonomy requires mastery or perfection of the skills underlying autonomy are ruled-out by a common-sense articulation of the empirical possibility/plausibility criterion for theoretical adequacy. One might think that we should begin a new critique by raising similar concerns about the additional criteria and skills outlined above. In order to adequately exercise autonomy, do we have to be able to:

---

[40] My only real objection is to the inclusion of the disposition to answer for one's self which has been developed in slightly different ways by Paul Benson and Andrea Westlund. But, I have included their view for completeness and my reasons for thinking this disposition is not a necessary component of a theory of autonomy are not relevant here—but basically it seems to me that their proposals boil down to much the same thing as self-respect.

imagine *all* possibilities that are open (or should be open) to us? possess perfect self-control? possess maximal self-respect?

But, I don't think this is the right way to proceed. It is both more fair and more efficient to take feminist and other thinkers at their word that they are attempting to develop a realistic, empirically achievable, model of autonomy (Govier 1993, 103) in much the same spirit as the arguments defending the practicability of autonomy surveyed in the previous section of this chapter. This would involve taking seriously those who stress that all the capacities and attitudes that are implicated in the exercise of autonomy are a matter of degree (Mackenzie 2014, 39, McLeod and Sherwin 2000, 266, Govier 1993, 105, Killmister 2013, 111), and that although in certain contexts it will be important to meet certain thresholds, perfection or mastery is never expected for people in general (Mackenzie 2014, 42). Taking these claims seriously allows us to by-pass a granular form of counter-argument in which worries about demandingness are raised for every capacity.

Even if we accept that the listed capacities should be taken as matters of degree and compatible with our common-sense beliefs about human limitations, there is still a question to be asked about the bigger picture. Adding so many capacities to the basic procedural model of autonomy (which is mostly interested in rationality and self-awareness) creates a model of autonomy that is at least highly *extensive* if not also highly *complex*. The reader will recall the distinction between *extensiveness* and *complexity* articulated in the introduction. By *extensiveness* I intend to refer to the number of things that are included within the idea of autonomy, in a way that is somewhat different than the regular philosophical idea of *extension*. Whereas normally the extension of a concept is all the things to which that concept applies, in this case I'm using *extensiveness* to refer to all the components of the concept. We might in this way say that the exercise of

autonomy has so many preconditions or elements that this potentially makes it problematically extensive. The exercise of autonomy becomes a burdensome task in which a wide range of skills and attitudes are invoked in the process of establishing one's values and living in accordance with them.

By *complexity*, on the other hand, I mean to pick out the relationships between the different elements of the concept.

*Extensiveness*s is the source of the problems I am raising in this chapter and the next, while *complexity* arises as a potential problem in Chapter 4. In collectively increasing the number of capacities involved in the exercise of autonomy, and therefore its *extensiveness*, the thinkers who proposed amendments to the basic model of autonomy in hopes of articulating what it would mean for autonomy to be "more embodied, more natural, more social, and more sensitive" (Govier 1993, 103), in a word more realistic, seem to have simultaneously provided the basis for renewed doubts about both the coherence and practicability of autonomy.

From the perspective of practicability, I would propose that the *extensiveness* that has been generated through recent theorizing about autonomy potentially makes it too difficult to exercise. That is, even if we accept that each of the skills and attitudes needs only be possessed to a certain degree, we might find the sheer number of simultaneous processes that are implicated overwhelming. Not only does each of the additional capacities and attitudes required for autonomy count as a further way to fail at being autonomous, but we may be inclined to think that no one could reasonably be expected to engage in such a multi-faceted undertaking. And yet, the common belief that people are in general able to exercise autonomy would suggest that not only is it possible to juggle so many capacities, but most people are able to do so.

The basic proceduralist model brought into question the autonomy of people whose cognition was significantly unformed or impaired (babies, people in comas, people with severe developmental delays or dementias) and those whose wills were alien to them (people under hypnotic suggestion, people with unwanted addictions or compulsions). Under the revised model it appears that the autonomy of an even broader range of people will be brought into doubt. For example, someone who is physically disabled or frail might lack *self-trust*, and might limit his activities in response to this. A fear of falling may result in decreased willingness to engage in exercise or errands. On the revised model, we would have reason for doubting this person's autonomy in a way that the simplified model did not. Or, a person with ADHD, whose executive functioning is compromised and therefore who might be described as facing a deficit in volitional or motivational competency, would also be seen as having limited autonomy even though this is a condition many people live with.

A second problem, which has to do with the effects of the added capacities on the coherence of autonomy as a concept will be discussed in depth in the next chapter. I noted in the introduction that a potential problem with the concept of autonomy is that it has too many meanings and is too complicated to put to work in making moral judgments (Arpaly 2002, 118-19). With so many additional features to consider and include in our characterization of autonomy, to critics it might seem that the term becomes even more elusive and difficult to employ. Thus, there are at least two possible critiques of an enriched understanding of autonomy: that it is too demanding, and that it further renders the concept of autonomy unintelligible. In the rest of this dissertation I will seek to address both of these problems in an effort to offer a more coherent and plausible conception of autonomy. In the next chapter, I will seek to address the coherence problem by putting forward a multi-level method for organizing the different concepts associated with

autonomy in to a coherent conceptual structure. In the fourth chapter, I will turn to the task of

showing that a richly developed account of the most fundamental level of autonomy is not overly

demanding despite the concerns raised above.

Chapter 3: Extensiveness Should not be Mistaken for Incoherence

The previous chapter concluded by raising two potential difficulties associated with the added nuance in theorizing about autonomy that has emerged in contemporary discourse. These problems, in my mind, were both the result of the way that these features added additional dimensions or preconditions to our understanding of the exercise of autonomy. One problem was whether, given that the resulting picture of autonomy is more extensive, it is plausible that people are generally capable of exercising autonomy to an adequate degree. That is, adding these additional features brings the practicability of autonomy into question. This issue will be the focus of subsequent chapters. Solving the second problem, having to do with the coherence of autonomy as a concept, will provide important resources for helping to defend the practicability of autonomy. The purpose of this chapter, therefore, is to address the potential excessive extensiveness of autonomy and its implications for coherence and intelligibility.

Having an extensive account of autonomy in itself is not necessarily undesirable. We might expect autonomy to be relatively conceptually complicated because of the rich and multi-faceted ways in which we use it (Dagger 2005, 196). Indeed, it was recognizing this attribute of autonomy that prompted Gerald Dworkin to resist "specify[ing its] necessary and sufficient conditions" (Dworkin 1989, 54). He worried that this would drain the richness from the idea, making it unable "to perform its theoretical role". My arguments in the next chapter will vindicate this intuition of Dworkin's[41], though I will rely on somewhat different considerations. That is, I believe autonomy-- as the sort of meta-capacity characterized at the end of the preceding chapter-- *is* and *should be* complicated in a way that would resist the specification of

---

[41] Which is also shared by Catriona Mackenzie (Mackenzie 2014, 18)

necessary and sufficient conditions. Indeed, my contention will be that it is only by taking all of the relevant elements of the concept into account that we can make theoretical sense, and satisfactorily precise use, of the idea. This chapter, however, will first contend with the charges that autonomy as a concept is overly or problematically extensive.

My approach to dispelling concerns about the coherence of autonomy will not pursue the strategy I think Christman had in mind when he lauded the usefulness of clarifying the concept of autonomy by seeking its conceptual core (Christman 1989, 4-5). I will not be attempting to unify all the disparate conceptions of autonomy under a single banner or core idea. Instead, I will argue that our sense of the coherence of autonomy as a concept can not only be salvaged, but enhanced, through a multi-level analysis of the exercise of autonomy. Such an analysis provides a system for fitting together many of the ideas that are invoked in discussions about autonomy and supplies the basis for an account of the relationships between these ideas.

I will begin this chapter by revisiting a prominent critique of the conceptual coherence of autonomy offered by Nomy Arpaly in *Unprincipled Virtue*. After articulating her critique and placing it within the context of her broader project, I will introduce my own account of the order that exists among the ideas that are associated with autonomy. I will propose and describe three levels of analysis, each of which I will illustrate by drawing on examples from contemporary philosophical work on autonomy. I will then use this analytical approach to show how, even if Arpaly is correct that references to a broad concept of autonomy are not useful for understanding the nature of moral responsibility, an account of the relationships between various ideas associated with autonomy-- and of their roles within a well-developed understanding of the exercise of autonomy-- can enhance our ability to make precise and coherent use of the concept. My intention is to convince the reader that it is possible to articulate a concept of autonomy that

can account for how and why the concept comes to implicate and integrate such an extensive number of other ideas. It is my contention that such a conception is useful for a wide range of philosophical purposes, including ones that complement or enrich Arpaly's own project.

Nomy Arpaly and the (In)coherence of Autonomy

In *Unprincipled Virtue*, Arpaly advances a theory of moral praise and blame-worthiness that is based on the quality (goodness of badness) of an agent's will and "the depth of their (moral) motivation" (Arpaly 2002, 115). Her account simultaneously attempts to treat cases that fall outside those usually considered in moral psychology (Arpaly 2002, 8) by not giving a privileged role to "conscious attitudes, deliberation, or… agent[s'] awareness of the moral significance of what [they are] doing" (Arpaly 2002, 117). Contrary to the shift she perceives in theorizing about moral responsibility, Arpaly resists introducing autonomy into her theory of moral responsibility (Arpaly 2002, 118). In her view, autonomy could only be as important to moral responsibility as other theorists often believe if it is "a precondition of praiseworthiness or blameworthiness, or…. it affects praiseworthiness or blameworthiness when all else is equal" (Arpaly 2002, 130). But, she notes, 'autonomy' is used in so many different ways that it's hard to say whether autonomy does, or even can, play this role.

In advancing her critique of autonomy, Arpaly therefore begins by cataloguing what she takes to be some of the multifarious senses in which 'autonomy' is used, what she calls the "varieties" of autonomy (Arpaly 2002, 117)[42]. Her list identifies at least eight—reluctantly nine—distinct

---

[42] Arpaly is not the only one to use this kind of strategy. Feinberg (1989), Khader (2011), and Gracia (2012) also employ approaches that identify and examine different kinds, types, or senses of autonomy in order to advance our understanding of autonomy or their positions regarding autonomy. One might even suggest that they use a common *thematic approach* in their analyses of autonomy.

things we might be referring to when we employ the word 'autonomy'. The meanings of autonomy Arpaly identifies are the following:

Agent autonomy, which she takes to refer to the relationship between people and their beliefs, desires, or values that makes them self-governing (Arpaly 2002, 118). The relationship is characterized differently by different theorists (for example, reflective endorsement versus alignment between first and second order desires), but the general idea is that this relationship establishes or reflects an agent's control or ownership over these elements of her mental life. As she develops her critique of autonomy's usefulness for the purposes of understanding moral responsibility, it becomes clear that for Arpaly *self-control* is the central aspect of agent autonomy (Arpaly 2002, 128, 132).

Personal efficacy, which she also calls material independence (Arpaly 2002, 119). According to Arpaly, this meaning of autonomy is the characteristic of being able to achieve one's ends in the world without the help of others. Wealth, certain skills (such as driving), knowledge, and physical strength are the means by which personal efficacy is attained.

Independence of mind or *psychological independence*, are at issue, Arpaly claims, in the cases of servile or submissive people (Arpaly 2002, 120). Such people seem to blindly accept and follow the views of their leaders—be they husbands, masters, political leaders, etc. And, this represents a kind of psychological diminishment, irrationality, or vice that is contrary to autonomy.

Normative autonomy, which Arpaly describes as the right to make one's "own decisions…free from paternalistic intervention" (Arpaly 2002, 121). This conception of autonomy is importantly social in that normative autonomy involves a kind of social respect or regard which not only

means that people are to be granted this kind of freedom, but also militates against interference by both political actors as well as other individual agents.

Authenticity, which according to Arpaly involves living and acting in accordance with one's deeply held values or "being true to oneself" (Arpaly 2002, 122). Those who espouse this conception of autonomy might see deeply held values as guiding behaviour both toward the selection of particular ends, as well as restricting our ability to bring ourselves to act in a way that is contrary to them. That is, on this meaning of autonomy we may be brought to act autonomously in this sense by our deeply held values despite ourselves (Arpaly 2002, 122).

Self-identification (Arpaly 2002, 125), according to Arpaly, is included in versions of authenticity on which one's sense of intra-personal coherence or integrity is central. This might seem to pull in the opposite direction of the previous idea that our deepest values could regulate our behaviour such that we may be autonomous despite ourselves. Instead, we might believe that authenticity involves the absence of such internal tensions or alienation from ourselves, and rather a sense of wholeness or harmony (Arpaly 2002, 124).

Heroic conceptions of autonomy, as described by Arpaly, refers to a category of versions of autonomy made up of idealized images of what perfectly autonomous people and their lives would be like. In the philosophical sphere, Arpaly presents us with examples such as Spinoza's vision of freedom, the contemplative life in Aristotle, and Nietzsche's Übermensch. But, in a more psychological vein she suggests we might also include the liberation aimed at by Freud and Jung (Arpaly 2002, 124). In some cases what makes these accounts heroic is that they involve the maximization of some of the other versions of autonomy Arpaly describes (such as rationality, self-control, self-efficacy). Heroic conceptions of autonomy are therefore by definition attainable only by the few.

Reasons responsiveness, Arpaly admits, is not widely believed to be equivalent to or the entirety of autonomy. But, she emphasises the important role that judgements about rationality play in judgements about autonomy. First, autonomy is meant to involve action in accordance with reasons as opposed to desires (Arpaly 2002, 126). Second, the process of endorsement that might be central to some conceptions of autonomy (like agent autonomy or authenticity) is thought to require *rational* deliberation. This close interconnection between reason autonomy makes it easy to believe that "[t]hings that impair our ability to respond to reasons… impair our autonomy" (Arpaly 2002, 126).

Moral responsibility is a final potential meaning of autonomy that Arpaly considers somewhat reticently. The purpose of her careful examination of the meanings of autonomy was to try to identify which of the things we call autonomy could most appropriately be deemed to be required for *moral responsibility* (Arpaly 2002, 118). Arpaly proposes that we might simply understand autonomy as "whatever it is that moral responsibility depends on" (Arpaly 2002, 139), though she ultimately does not accept this conclusion.

In reviewing these possible meanings of autonomy, Arpaly encounters several problems. A preliminary set of her concerns relates to the various ways in which she thinks at least some of these meanings are problematic in their own right. For instance, heroic varieties of autonomy and conceptions of autonomy that would demand complete and unbiased information threaten to place autonomy outside the abilities of the average person (Arpaly 2002, 121, 124). And, the feelings of alienation given extensive treatment in self-identification forms of autonomy are a poor guide to establishing which self-defining attitudes are reflective of the real self. But, beyond these concerns about the individual meanings of autonomy, Arpaly's main point of contention with the use of autonomy is the apparent lack of connection between the different meanings. On

her description lacking (or possessing) some forms of autonomy, reveals little about whether one is deficient in other forms (e.g. Arpaly 2002, 120). Indeed, she observes, some meanings of autonomy seem to lead to conflicting conclusions and judgements (Arpaly 2002, 128-29). Autonomy as a concept, Arpaly concludes, is therefore at best "too elusive and complex" to be of use in clarifying other terms (Arpaly 2002, 118). At worst it is incoherent.

For Arpaly, there is considerable intellectual danger in failing to distinguish between the various meanings of autonomy she has identified in the process of attempting to understand moral responsibility through autonomy. First, some senses of autonomy, such as autonomy as a right or as a heroic ideal, have little or no relevance to moral responsibility (Arpaly 2002, 130). So, there is some need to winnow down the range of types of autonomy considered to only those that are relevant to responsibility if we are to (as Levy says (2006, 429)) "make progress" on the question of how autonomy and responsibility are related. If the discussion is not restricted to only relevant autonomy types, then people are liable to be influenced by their intuitions about irrelevant senses of autonomy (Arpaly 2002, 127), which then distorts what we take to be true about moral responsibility (Arpaly 2002, 118). And, appealing to irrelevant or insufficiently specified conceptions of autonomy might prejudice theorists and readers for or against particular notions of moral responsibility, or distract them from the relevant considerations. Since she believes that a broader concept of autonomy yields unreliable and possibly irrelevant intuitions, Arpaly thinks it is better to use the meanings she has isolated as independent concepts for the purpose of theorizing about moral responsibility.

While Arpaly may well be right that no direct reference to autonomy is needed to understand moral responsibility, this does not mean that (i) there is as little unity among the ideas associated with autonomy as she suggests or (ii) that an appropriately developed complex concept of

autonomy can be of no use in understanding the relationship between moral responsibility and autonomy. In what follows I will advance a way of analysing the exercise of autonomy that I think allows us to bring together many of the ideas associated with autonomy into a coherent picture. And, I will show how this way of understanding autonomy provides new resources for giving precise and nuanced responses to questions that complement those Arpaly is asking about the link between autonomy and moral responsibility

Contending with Extensiveness

The approach to managing the conceptual extensiveness of autonomy I will advance builds on the picture of autonomy that we were left with at the end of the previous chapter. That model depicted autonomy as a capacity or ability that was made possible by a host of underlying factors including both capacities and attitudes. Taking inspiration from this idea, I propose that we should come to grips with the conceptual extensiveness of autonomy by focusing our attention on the levels at which autonomy can be analysed. In advancing this view, I am proposing a hierarchical way of organizing the extensive number of elements associated with autonomy. But, this is a hierarchy of a very specific sort. There are other approaches to examining the different aspects or kinds of autonomy that also involve some kind of hierarchical organization. For example, in arguing that reduced autonomy cannot explain adaptive preferences, Khader first divides autonomy into substantive and procedural categories, and then divides these into various types and subtypes (Khader 2011, 78-92, 97-103). Feinberg also divides autonomy up into types (a capacity, an ideal, a condition, a right) and then further subdivides his conception of autonomy as a condition into a range of further virtues he thinks autonomous people will have. What's important to recognize about the kind of hierarchy that exists within these views is that it is static and taxonomic in nature, merely a successive subdividing of categories.

The kind of hierarchical organization that I will be advancing in the rest of the chapter is different in several respects. First, it divides the concept of autonomy into two main parts. One part has to do with our judgements relating to autonomous people, the other part has to do with the exercise of autonomy itself. Importantly, I see the elements of each part as really being parts of autonomy in a non-classical[43] sense of parthood that might be accepted under something like Kit Fine's pluralist theory of parthood, which allows for compositional operations to take a wide range of forms and be based on a variety of principles (Fine 2010, 573). I further divide the exercise of autonomy itself into two hierarchically related sets of parts. The elements of the lowest level serve as parts in the form of constituents or precursors, of the other, while the elements of both qualify as parts (though sometimes in a mediated sense (Fine 2010, 577)) of autonomy as a capacity that is exercised. Thus, within the exercise of autonomy the relationship between different kinds of parts is something like a building relationship. The relationship between the exercise of autonomy and our judgements relating to autonomous people, however, is of a very different sort. I envision the judgements aspect of autonomy to be hierarchically above the exercise aspect of autonomy, with the exercise aspect serving to elucidate, justify, and explicate elements at the level of judgements.

The reader should be able to see that the kind of hierarchy I am proposing is significantly different than the kind of repeated sub-division of categories that would be employed in a more taxonomic approach to organizing the various ideas associated with autonomy. The hierarchy I am employing takes the items in the hierarchy as interacting elements of a system of concepts

---

[43] In a mereological sense.

through which we understand autonomy. In the paragraphs below I will more fully articulate the

rich interrelationships I think exist between the different levels of the hierarchy.

Layer Cake

In this sub-section I will describe three levels of analysis-- which I will label the *meta-*, *meso-*,

and *micro-*levels-- at which I intend to conceptualize autonomy. The meta-analytic level captures

the entirety of the part I described in the previous section dealing with our ideas about what

autonomous people are like. The other two levels comprise the second part described above

relating to the exercise of autonomy. Within an analysis of the exercise of autonomy, the meso-

level addresses the general functions that are implicated in the exercise of autonomy, while the

focus of the micro-analytic level is the means by which the functions at the meso-level are, or

can be, achieved. We have already encountered at least some of the elements of this lowest level,

namely capacities such as rational deliberation and attitudes such as self-trust. I will discuss each

of these levels in turn drawing on existing views to illustrate what autonomy looks like at each

level.

   i.   The meta-level of analysis

This highest level of analysis for conceptualizing autonomy is populated by ideas and normative

judgements about what autonomous people are like. This, I would suggest, is the major

difference between this level and the subsequent two. While at the meta-level the topic of

discussion is more centrally autonomous *people,* the exercise of autonomy itself is the focus of

the lower two levels. Certainly, the autonomous person is in the picture with regards to the other

two levels as they are the ones exercising autonomy. For example, we might say that "to be

autonomous a person needs to be able to perform functions X, Y, and Z." The ability to perform

the enumerated functions to a satisfactory degree is a property of the autonomous person, and

this ability is what makes their exercise of autonomy adequate. If someone is autonomous, then we can reasonably believe them to be capable of the autonomy functions, but we might also believe that certain traits or characteristics issue from the condition of being autonomous. The emphasis in the meta-autonomy level is on these latter kinds of claims. For instance, this is the level at which virtues presumably associated with autonomy are pitched. Feinberg proposes that autonomous people will have integrity and initiative (Feinberg 1989, 39, 41), but these are properties of people not of the capacity. Presumably these virtues are a consequence of exercising the functions that are required for autonomy, but are neither equivalent to it nor constituents of it. Normative conceptions of autonomy dealing with the kind of social, political, and legal status autonomous people should enjoy would also be pitched at this level. The central concern of these types of views is what rights people have or what treatment they deserve from other people and from social institutions when they are deemed to be autonomous.

An example of a discussion of the nature of autonomy that I see as being offered almost entirely at this level is Feinberg's essay "Autonomy". Feinberg identifies "four closely related meanings" of autonomy (Feinberg 1989, 28). These are autonomy as a capacity, a condition, an ideal, and a right. In the essay, which is a reprint of Chapter 18 of *Harm to Self*, he provides an account of each of these meanings of autonomy as they are applied to people. For instance, autonomy as a condition is further broken into various virtues that autonomous people supposedly possess. In this section, I will not articulate his view in detail. I will only highlight some of the more surprising places where his analysis remains at the meta-level.

The first aspect of Feinberg's piece that is worth noticing is how he discusses autonomy as a capacity. Though I have to this point only provided a very general sense of the structure of the scheme I am proposing for dealing with the conceptual extensiveness of autonomy, my

comments at the beginning of this subsection (and in the introduction) contain the suggestion that looking at autonomy from the perspective of capacities is characteristic of the micro-level of analysis. So, from what I've said so far it would be reasonable to expect that Feinberg's account also has something to say about the lowest level of analysis.

To be fair, although his section on autonomy as a capacity says very little about what autonomy is like as a capacity or what capacities might be involved in autonomy, there are two moments where Feinberg's treatment of autonomy as a capacity is similar to what might be expected at my micro-level of analysis. First, he does say that the capacity for self-government (by which he means autonomy) "is determined by the ability to make rational choices" (Feinberg 1989, 28), though his reference to the "capacities relevant to self-government" later in the section suggests that either the ability to make rational choices involves many capacities, or that it may not be the only capacity required for autonomy (Feinberg 1989, 30). But, this is not given any further attention. Second, he claims that the capacity for autonomy is possessed by most people as a matter of degree, but is not possessed by people with various mental defects (infants, people in comas), non-human animals, plants, or inanimate objects (Feinberg 1989, 28-30).

It is significant for our purposes that, aside from these few points, Feinberg spends most of the section discussing the rights someone comes to possess when they are deemed to achieve the threshold for competence with respect to autonomy in society, and the kind of freedom they are allowed. People above this threshold are deemed to be equally qualified to govern themselves, and being so qualified an autonomous person is, on this view, free to "rule himself badly, unwisely, [or] only partially" as much as prudently or sagely (Feinberg 1989, 30). Feinberg's exposition on this topic says little about how autonomy is achieved through the exercise of capacities, focusing instead on what life is like for those who meet the threshold, and how they

should be viewed socially. As such, his account of autonomy as a capacity is largely irrelevant to the micro-level analysis I have in mind, and is much more reflective of the kinds of considerations that are of interest at the meta-level.

The second part of Feinberg's essay that I want to attend to rather carefully is the way he discusses authenticity in the form of *self-selection*. At least superficially, self-selection sounds like it might be related to the meso-level function of *self-definition* proposed by Diana Meyers, which will be discussed in the next section (Meyers 1989, 43). Indeed, Meyers claims that they are the same thing (Meyers 1989, 44). Some of Feinberg's description of self-selection, which on his view is a virtue associated with autonomy rather than part of how one becomes autonomous, does emphasise the role of rationality in scrutinizing and modifying one's opinions and tastes. In this moment his section on self-selection does sound more like what I will argue is involved at lower levels of analysis.

However, much of the section is taken up with a discussion of the properties that a person who self-selects will and will not have. We learn that such a person will not be "merely the mouthpiece of other persons or forces", will not be "a[n] habitual and uncritical conformist… [or a] manipulated consumer" (Feinberg 1989, 32). Furthermore, a self-selecting person, on Feinberg's account, will not possess ideals implanted by his elders that are impervious to rational criticism. On this view, an autonomous person also consumes products that he can afford and that meet his personal needs, and chooses a lifestyle that reflects his personality (Feinberg 1989, 32-3). These remarks take up the better part of the discussion of what could potentially count as an autonomy function. This section is dedicated not to articulating how this function supports autonomy, but rather to anticipating the kinds of behaviours that might be associated with it in

autonomous people. Because of these features this discussion is also pitched at the meta-level even though the topic could be given more precise treatments at lower levels of analysis.

While these examples represent only part of Feinberg's account of autonomy as it appears in "Autonomy", they do illustrate how his discussion of autonomy remains at the level of analysis that considers the nature of autonomous people, rather than the nature of autonomy as a capacity, at important junctures. On other levels of analysis, as we shall see in the next two sub-sections, capacities and self-selection/self-definition are addressed much differently.

ii.   The meso-level of analysis

Several accounts of autonomy that I would argue possess a meso-level dimension have been advanced in contemporary work on autonomy, though they were not necessarily developed with this intention. As I have expressed above, on my view the meso-level of analysis of autonomy as a concept is comprised of the functions that people who exercise autonomy must be capable of. Certainly, there are ways of describing these functions to make them about people. For instance, if *self-definition* is a function that is part of autonomy then autonomous people must be self-defining or capable of a certain satisfactory degree of self-definition. The difference that I believe exists between the meta-level and meso-level analysis of autonomy has not only to do with whether it is autonomy or autonomous people being described, but also whether (a) the capability, virtue, etc. under consideration is involved in exercising autonomy, or (b) it is being proposed as the consequence of the exercise of autonomy or a judgement about those who exercise it. The consequences or judgements are, on my classification, properly the subject of the meta-level of analysis, whereas the exercise of autonomy is of interest at the meso- and the micro-level of analysis.

The difference between the meso-level of analysis considered in this sub-section, and the micro-level of analysis to be considered in the next sub-section, is somewhat more difficult to draw. One reason for this is that meso-level analyses are often constructed with reference to features of the micro-level. That is, while a number of writers have identified sets of mid-level functions that form the basis of autonomy, they also tend to recognize that these still complex functions are the product of the confluence of still further factors—underlying skills, attitudes, etc. So, the functions that I take to be characteristic of the meso-level of analysis are themselves also not unitary or monolithic in nature. But, they are at least more granular than just the broad concept of autonomy in that they identify some aspects of what choosing/discovering one's values and living according to them requires or involves for individuals. The general description of autonomy we are working with suggests two major categories of functions that are part of autonomy— those that are related to authenticating values, which might be cashed out in different ways depending on whether one understands reasons and the good to be objective/external or subjective/internal[44]— and those that are related to living according to one's values.

To illustrate what a meso-level account of autonomy is like, I will focus on a recent book-length treatment of the concept of autonomy by Suzy Killmister. She advances a four-dimensional theory of autonomy consisting of self-definition, self-realisation, self-unification, and self-constitution. Her model will also be compared and contrasted with meso-level accounts offered by Diana Meyers and Catriona Mackenzie. Examining these three versions of meso-level accounts will reveal that, while there are some shared ideas and intuitions about what autonomy

---

[44] As noted in the introduction, the three-level analysis and in particular my way of conceiving of the exercise of autonomy might be adopted by both camps since the processes of choosing and adopting reasons rely on similar sets of abilities. The primary difference would likely be which functions are used most, or are most important.

involves, the application of labels to ranges and sub-sets of functions comprising autonomy (for indeed there may be different layers even within the meso-level of analysis) is still unsettled.

The first element of Killmister's theory is *self-definition*. According to Killmister, self-definition relates to whether someone is self-governing (for her a synonym for autonomy) with respect to their personal identity (Killmister 2018, 19). As such, she takes self-definition to describe the degree of coherence between an individual's personal identity and their self-defining attitudes (Killmister 2018, 33). She describes personal identity as "the cluster of features that make [someone] who she is, and that connect her identity over time". These features include, but are not limited to, self-defining attitudes. Self-definition, according to Killmister, means establishing and coordinating between one's self-defining attitudes—beliefs, values, and goals (Killmister 2018, 20). These attitudes, which are grounded in commitments, will demand to be treated in particular ways. She thinks they will generate normative requirements or will be the kinds of things about which we have self-governing policies that establish rules about what self-defining attitudes we can become committed to and how we treat these attitudes (Killmister 2018, 22, 29). For Killmister, autonomy requires consistency between self-defining attitudes and self-governing policies so that they can reside coherently within a person's identity (Killmister 2018, 31). But, on her view, self-defining attitudes themselves play a role in establishing the necessary consistency and alignment as they at least partly dictate what new self-defining attitudes are taken on, and at the same time allow for some readjustment in terms how weighty and central they are within the "web" of self-defining attitudes that shape one's identity as new self-defining attitudes are adopted (Killmister 2018, 31).

The second dimension of autonomy on Killmister's account is *self-realisation*. On her view, self-realization is closely connected with practical agency, which is the reaction (affective,

interpretive, deliberative) of an agent in response to her environment (Killmister 2018, 35). To be autonomous, Killmister thinks that people need to be both internally and externally self-realising. She contends that subdividing self-realisation into these two parts helps distinguish between separate processes that are involved, which are open to different kinds of failures. Internal self-realisation has to do with setting intentions as the result of practical deliberation. In successfully exercising autonomy people's intentions will reflect the balance of reasons they see themselves as having for and against a given option. This means that, for Killmister, autonomous people will not develop intentions that the balance of reasons points against (which might be seen as succumbing to temptation), or act on reasons that they have excluded from their deliberations altogether (Killmister 2018, 37).

External self-realisation concerns the "connection between intending and acting" (Killmister 2018, 40). In autonomous people, external self-realisation involves a smooth transition between intention and action. Our intentions might produce in us commitments about *why or how* an action is to be performed, or what counts as success (Killmister 2018, 40). Autonomous actions will satisfy these intentional commitments, whereas failure on one of these points is reflective of reduced autonomy. There are therefore a variety of ways in which people can fail to be externally self-realising. These include instances in which there is no transition from intention to action, or when the transition is imperfect. When people are physically restrained by others from, or when they simply fail entirely on their own to, follow through on their intentions in action they may be lacking entirely on this dimension of autonomy. According to Killmister, external self-realisation may partly fail to satisfy the demands of autonomy when one is deceived about the nature of one's actions or when one acts for reasons other than the ones that one had initially planned to act on.

Killmister's third domain of autonomy is *self-unification*. Its primary feature is the relationship between what emerges from self-realisation and what emerges from self-definition. Without this domain, it is possible that people could qualify as autonomous while their commitments undertaken in processes of self-definition are entirely different from those formed in processes of self-realisation. In Killmister's words, "an agent might be committed to acting a certain way, without being similarly committed to deliberating in a certain way" (Killmister 2018, 63). But, according to Killmister, for someone's self-realization and self-definition to count as part of an adequate exercise of autonomy, commitments emerging from both domains must be coherent and consistent. For example, the commitments attached to one's self-defining attitudes will constrain practical deliberation (or even our affective responses) resulting in coherence between the two sets of commitments (Killmister 2018, 59). For autonomous people, there will even be alignment between unreflective action and self-defining attitudes (Killmister 2018, 64). Killmister's discussion reveals that there are many ways that someone can fail to fulfill this aspect of autonomy. We might fail to be self-unified when we fail to uphold our self-defining attitudes in practical deliberation (Killmister 2018, 58), when we commit ourselves to not experiencing particular affective states (like jealousy or fear (Killmister 2018, 61)), and yet feel them despite ourselves, or when our knee-jerk reactions fail to uphold the commitments associated with our self-defining attitudes (Killmister 2018, 64).

Killmister's final dimension of autonomy is *self-constitution*. She contends that this is the dimension of autonomy that is required in order to be autonomous in a global, rather than local, sense. Self-constitution considers the extent to which agents are even involved in the kinds of activities characteristic of self-definition and self-realisation—the extent to which they adopt self-defining attitudes, or "engag[e] in practical deliberation and intention formation" (Killmister

2018, 69). In order to be self-constituting, Killmister thinks people have to meet three conditions. First, they must have adopted self-defining attitudes, and therefore have a personal identity that contains normative commitments. Second, they must be "able and willing to exercise practical agency", which means having the capacity to and, when appropriate, engaging in practical deliberation and intention formation, as well as taking the outcomes of these processes seriously as constraints on behaviour (Killmister 2018, 71). Finally, self-constitution relies on people's ability to unify their personal and practical identities.

On Killmister's view, self-constitution plays an important role in establishing the threshold of autonomy that people must meet if they are to be granted the social consideration and respect we tend to believe is owed to autonomous people (Killmister 2018, 70-1). But, she thinks that most people, with the exception of young infants and people with severe cognitive disabilities, will be able to meet all three criteria of self-constitution. Some boundary cases include people who have been gaslighted and therefore do not trust their practical deliberations, and those with advanced Alzheimer's disease. The former would fail to meet the second condition since they would not take the outcomes of their practical deliberation seriously, and the latter would fail the third condition because their disease would prevent them from unifying their self-defining attitudes with practical deliberation (Killmister 2018, 72-3).

Killmister's view is different from the aspects of Diana Meyers's *Self, Society, and Personal Choice* and Catriona Mackenzie's "Three Dimensions of Autonomy" that I read as meso-level analyses in perhaps one crucial respect. Perhaps because of Killmister's focus on commitments that arise in what she sees as two aspects of the self, personal identity and practical agency (Killmister 2018, 11-12), her view is less explicitly connected to the micro-level analysis. So, whereas Meyers and Mackenzie stress that the dimensions or components of autonomy are

themselves reliant on further skills or conditions (Meyers 1989, 53, 59, Mackenzie 2014, 18)

Killmister's account does not pay these much attention. Killmister's model still suggests these

relationships, however. For instance, in her discussion of the case of gaslighting mentioned

above, she notes that trust (presumably self-trust) is required for taking one's deliberative

processes seriously, which undermines both self-realisation and self-constitution (Killmister

2018, 72). So, the link between the meso- and micro-levels of analysis may still exist on

Killmister's view, even if the way in which she articulates the functions at the meso-level does

not make it explicit.

There is also important overlap between Killmister's model and those of Meyers and Mackenzie.

In the aspect of Meyers's account of autonomy that contains a meso-level analysis, at least three

functions are named: self-discovery, self-direction, and self-definition. Self-discovery and self-

definition are ways in which we come to have or know our personal identities. Self-discovery is

the processes of coming to know one's values and beliefs. According to Meyers self-discovery is

problematic on its own since it suggests that people have some true, trans-social core self

(Meyers 1989, 44-5). Meyers thinks that a full account of autonomy must also include self-

definition, which is the process by which people come to "conceive and institute changes in their

true selves" (Meyers 1989, 43). Together it would seem that Meyers's self-discovery and self-

definition overlap with Killmister's self-definition and at least the first condition of her self-

constitution.

Self-direction on Meyers's view takes two forms. One is local (or in her words episodic), and the

other is global (or in her words programmatic). Episodic self-direction involves the exercise of

practical rationality and as such seems to capture much the same functional territory as the two

aspects of Killmister's self-realisation. Meyers's programmatic self-direction, which involves

thinking in a global way about the kind of person one would like to be and how to realise this (Meyers 1989, 48-49), seems to range over aspects of each of Killmister's dimensions, involving adopting self-defining attitudes, as well as making and carrying out intentions for pursuing them in a way that is unified. Indeed, Killmister's self-unification dimension may be reflected in the aspects of Meyers's view that stress the importance of personal integration (Meyers 1989, 62).[45]

Mackenzie also proposes three axes of autonomy, two of which I think qualify as meso-level functions. These are self-governance and self-authorization. Mackenzie's self-governance is most like what appears in the accounts already considered. It consists of all "the skills and capacities necessary to make choices and enact decisions that express, or cohere with, one's reflectively constituted…practical identity" (Mackenzie 2014, 19). Once unpacked this axis of autonomy is revealed to contain quite a range of processes and conditions that appear to overlap with much of what Killmister has divided into self-definition, self-realisation, and self-unification. Meanwhile, self-authorization does not necessarily have a straightforward analogue on either Meyers's or Killmister's account. It involves a sense that one has *normative authority* over one's life. This involves "regarding oneself as authorized to exercise practical control over one's life, to determine one's own reasons for action, and to define one's own values" (Mackenzie 2014, 19). While Meyers contends that autonomous people will be actively engaged in shaping their own lives (Meyers 1989, 47) and Killmister postulates that the self-constitution aspect of autonomy requires that people take on self-defining attitudes and that they be willing and able to engage in practical deliberation, neither quite captures the self-regard or self-understanding that is suggested as essential to autonomy under the function of self-

---

[45] This serves as a hint that in trying to establish what the functions underlying autonomy actually are we should not become fixated on proposals that take the form of "self-x" that we ignore relevant proposals.

authorization—with the exception perhaps of instances such as Killmister's indirect acknowledgement of the importance of self-trust for the exercise of certain autonomy functions.

Finally, Mackenzie's third axis reveals an important distinction that can exist between functions. There are some functions like Killmister's self-definition, and Mackenzie's self-government that make reference only to internal states and processes within agents (Killmister 2018, 39, Mackenzie 2014, 19). But other functions, including Meyers's self-direction and Killmister's external self-realisation, have very much to do with the relationship between an agent and her environment. Mackenzie's third axis, self-determination, is primarily concerned with the ways that one's environment shapes one's ability to enact the internal aspects of autonomy in the world. This axis of autonomy articulates the freedoms and opportunities that make the exercise of other autonomy functions, and therefore autonomy, possible (Mackenzie 2014, 18). But, it is not itself a function and is therefore not an immediate consideration at the meso-level of analysis.[46]

In these examples of meso-level analyses, a number of functions that underpin the exercise of autonomy— including self-direction, self-governance, self-authorization, self-discovery, self-definition, self-unification, and self-constitution— have been proposed. However, these should not be taken as exhaustive of the potential elements of analysis at this level.[47] What the enumeration of these functions, and the considerable consistency between the views presented, suggests is that there is broad agreement in our intuitions about what kinds of things people must

---

[46] Though social conditions could be very relevant to establishing how much potential someone has to express particular functions and related questions.

[47] The descriptions I have given of them should also not be understood to be the only ways of articulating what each of the functions is supposed to capture. Meyers, Killmister, and Mackenzie's accounts are all internalist/subjectivist about what is important (though it is worth noting that Meyers's internalism is meant to be constrained by the limits of objective morality). The functions involved in authenticating values and directing one's life/acts according to them might be articulated in different terms by theorists of a more externalist/objectivist stripe. Thanks to R. Myers for this point.

be able to do in order to live according to their authentic values. In the next section, I will discuss how these mid-level components of autonomy themselves are constituted by still more basic elements.

iii. *The micro-level of analysis*

The process of describing this final level of analysis has already been initiated in earlier sections. In the previous sub-section, I proposed that the difference between the meso- and micro-levels of analysis was that the former are to be understood as produced by the exercise of the latter which consist of skills, attitudes, and other factors. This relationship is explicitly recognized by Catriona Mackenzie when she specifies that each of her axes of autonomy rests on still further conditions (Mackenzie 2014, 22) and implied by Diana Meyers in the moments where she discusses the skills people employ in carrying out processes of introspection and self-determination, i.e.; the skills that are needed to possess autonomy competency (Meyers 1989, 53). As to the kinds of things we would expect to attend to at the micro-level of analysis a preliminary discussion already appears near the end of the previous chapter. In the section "A New Common-Sense Concern", I enumerated a number of different elements that have been proposed as forming the basis for autonomy. I will only briefly rehearse some of these below. The brevity of my treatment of this level of analysis at this stage is partly because I have already offered a preliminary description of this level in the previous chapter, and also because this level will be the central topic of the remaining chapters.

We might begin to populate the micro-level by pointing to the ways in which procedural and substantive accounts of autonomy have contributed to the elaboration our understanding of what is involved in autonomy. Mackenzie proposes that "the kernel of what is taken to be the unitary concept of autonomy is the notion of self-governance", which has led to a focus on the skills

needed to make self-defining attitudes one's own and the skills needed to translate self-defining attitudes into action effectively (Mackenzie 2014, 17-8). These are the authenticity and competency conditions, respectively. Authenticity skills might include capacities such as introspection, rationality, self-awareness and imagination, whereas competency skills might include self-control, resolve, deliberative skills, motivational effectiveness, and decisiveness.[48] In addition to these capacities a number of substantive pre-requisites to autonomy have been proposed. These have included sanity as well as attitudes such as self-trust, self-confidence, self-worth, self-esteem, and self-respect.

In my view, these kinds of skills, attitudes, and predispositions are some of the means by which meso-level functions are achieved. A full articulation of the relationship between the means of the micro-level and the functions at the meso-level will be reserved for the next chapters. But, we can begin to appreciate the general picture by considering an example. Take, for instance, self-definition. While it is true that this function is articulated in slightly different ways by different authors, we can relatively easily imagine which kinds of means would be required for it. Self-definition is part of establishing authenticity, and therefore is a function through which people's values become genuinely their own. To be self-defining we might expect that people need to be able to do at least some of the following:

- Know what values they already have and make judgements about whether these are the values they want

- Know about the world so that they can make an assessment of the processes by which they came to have their values, and so that they can learn about alternative values

---

[48] See sections "Authenticity Beyond Rationality and Self-awareness" and "Competency Skills" in the previous chapter for a more extended discussion.

(Killmister 2013, 112). Knowing about the world would also contribute to their ability to make judgements about which self-defining attitudes they favour given the way the world is or could be.

- Make assessments of and ultimately judgements about adopting or rejecting the options they entertain

Self-definition would therefore reasonably be expected to be achieved through capacities involved in self-knowledge like introspection (Meyers 1989, 46), as well as abilities relating to obtaining information about and interpreting the world, and skills relating to making judgements like critical reflection, imagination, and detecting and reconciling conflicts among components of the self (Meyers 1989, 53). The operation of these capacities might additionally be supported by feelings of self-worth and self-trust that provide the basis for our beliefs that coming to one's own idea of what is good and desirable is a worthwhile task, and that we can rely on ourselves to come to reasonable and appropriate conclusions (Mackenzie 2008, 525, McLeod and Sherwin 2000, 262).

While this example might help illustrate how the micro- and meso-levels of analysis are related, in the next chapter I will argue that the structure of the relationship is actually considerably more complex. But, this discussion should be enough to ground my position on the questions with which this chapter began. Namely, can a multi-level approach to autonomy succeed in managing the extensiveness of the concept of autonomy in a way that provides useful insights for the resolution of various philosophical problems? I will defend an affirmative answer to this question in the next (and final) section of this chapter.

How a Multi-Level Approach Helps

There are some things that the system of analysis I have proposed above clearly does not do. For example, my three-level analysis provided above does not establish a single core idea from which all possible conceptions of autonomy can be traced. Nor does it articulate a finite set of necessary and sufficient conditions for autonomy as a concept. In this sense, I am operating in much the same spirit as Catriona Mackenzie in conceiving of autonomy as a multidimensional concept, which does not support the existence of a single core idea and makes difficult the task of fixing a unified set of necessary and sufficient conditions (Mackenzie 2014, 16-7). In what follows, I suspect that the meso-level functions will come closest to playing the role of necessary, and collectively perhaps even sufficient, conditions for autonomy. But, to foreshadow, in the final picture of the exercise of autonomy I will present it will not be clear that much has been gained by knowing these conditions since to focus on them would be to lose sight of the considerable variability in the ways they can be realized. It is also worth noting that I do not endorse any particular set of functions as the correct ones.

Given these observations, it should be clear that what I have advanced is more of a system for understanding what autonomy involves, and for analysing the demands of particular conceptions of autonomy with the support of our common-sense intuitions about human psychology. These features are not so much weaknesses in my account as boundaries that delineate my approach from others. As I see it, a multi-level analysis of the concept of autonomy makes two important contributions to our understanding and use of the concept that are relevant to the critique Arpaly leveled against it. The first is the way it integrates many ideas associated with autonomy into a cohesive system in which relationships between the elements are clearly specified. The second is the conceptual resources it provides for extending the analysis of the relationship between autonomy and moral responsibility in a way that I think supplements Arpaly's stance on the

question. These resources, however, can also be applied to a wide range of conceptual questions about autonomy and form the basis on which we can give them precise evaluations.

In regard to the first contribution, my multi-level analysis proposes that many of the things that Arpaly identified as *kinds* of autonomy should instead be understood as *aspects of* autonomy. If we accept this shift in perspective, then we can see that-- contrary to the lack of correlation Arpaly found between different "varieties" of autonomy— there are at least sometimes good grounds for inferring or anticipating deficits in some aspects of autonomy based on deficits in others.[49] The account I will give of how the different levels of analysis are connected to each other, and how detrimental some deficits should be expected to be in the rest of the system, will be the central topic of the remaining chapters. To hint at what's to come, we can see how different aspects of autonomy might vary together by tracing the relationships between them and being cognisant of the roles they play in the exercise of autonomy. For example, someone might experience a deficit in a meso-level function if they also experience a deficit in a micro-level means that is a pre-cursor of that function. Or, someone might experience a deficit in several functions if these functions share a precursor that is also impaired. Placing the different ideas associated with autonomy into a system of relationships based on their role in the exercise of autonomy shows the concept to be more coherent than Arpaly seems to give it credit for.

With respect to the second contribution that I take the multi-level analysis to make relating to Arpaly's investigation into the relationship between autonomy and moral responsibility, it's

---

[49] There might be some concern that I am simply begging the question against Arpaly here since such a shift is not obviously defensible on her grounds. I would draw the attention of a reader with this kind of concern to Arpaly's discussion about the relationship between agent-autonomy and moral responsibility near the end of Chapter 4 of *Unprincipled Virtue*. Here, she claims that the two are connected because they both demand higher faculties, especially in the form of reflective abilities, as psychological preconditions (Arpaly 2002, 147). All I'm proposing to do here is employ as similar kind of analysis to the ideas associated with the broader concept of autonomy. Certainly, I can't address all of Arpaly's concerns in the process, but I hope the reader will agree I can make considerable progress. I'll discuss this more in the following pages.

useful to note two things. First, Arpaly argued that, of the eight or nine different meanings of autonomy she identified, responsiveness to reasons, authenticity, and to some degree agent-autonomy (i.e. self-control) are relevant to moral responsibility (Arpaly 2002, 131). From her claim that moral praise and blame-worthiness depend on agents' abilities to be motivated by moral reasons and the depth to which they care about moral reasons, she takes it to follow that moral responsibility depends on responsiveness to (moral) reasons and authenticity. Her view is that being motivated by moral reasons is a function of reasons responsiveness, while one's level of caring is related to authenticity (Arpaly 2002, 131). She also thinks there is an important relationship between what she calls agent-autonomy/self-control and moral responsibility. Although she thinks it is incorrect to believe that agent-autonomy grounds moral responsibility, she proposes that they rely on shared psychological preconditions, such as reflective abilities (Arpaly 2002, 147). Second, as we have begun to see, and will continue to see in subsequent discussion, all three of the varieties of autonomy that Arpaly thinks are relevant to moral responsibility are candidates for inclusion at either the meso- (self-control and authenticity) or micro-levels (reasons responsiveness[50]) of analysis.

My use of the multi-level analysis I have advanced opens the possibility of giving a different kind of analysis of the way the ideas associated with autonomy are related to autonomy than the one Arpaly gives. Much like her own assessment of the relationship between self-control and moral responsibility, my approach suggests that autonomy and responsibility depend on some of the same functions and capacities, but are only partly overlapping as concepts and certainly not equivalent (Feinberg, 1989, 43). This, I think rightly, conflicts with those who have treated

---

[50] I place reasons responsiveness at this level because what Arpaly means by the term is "the ability to respond to reasons" (Arpaly 2002, 125). So, while other interpretations of reasons responsiveness might see it as something more like a disposition or trait, on her view it's something more like an ability.

autonomy and responsibility as equivalent in philosophical discourse. Arpaly's own arguments suggest that autonomy and responsibility overlap in requiring the capacity for (what she calls) reason responsiveness and authenticity. Given their shared interest in authenticity, autonomy and moral responsibility might also both be grounded in functions like self-direction and capacities like introspection and critical self-reflection, which, based on my analysis, are some of the capacities underlying authenticity itself. There is also, as she notes, a point of overlap between self-control and moral responsibility based on their shared psychological preconditions. Although Arpaly's question about the relationship between autonomy and moral responsibility was articulated without reference to a multi-level system of analysis, we can see that through attention to the lower levels of analysis encouraged by such an approach we can provide a very precise account of the relationship between autonomy and moral responsibility despite retaining an extensive conception of autonomy, and without assuming that either is a prerequisite for the other.

Second, we might also use the multi-level analysis to answer the question of why people might have thought that autonomy was the grounds for moral responsibility (or equivalent to it) in the first place. Our analysis above suggests that, in general, given the similar roles that reasons responsiveness and authenticity play in both autonomy and moral responsibility our intuitions that track these aspects of a case are likely be led to similar conclusions about both concepts. That is, all else being equal, I expect a deficit in reasons responsiveness (or authenticity) would likely lead us to similar kinds of doubts about a person's autonomy and praiseworthiness (see Arpaly 2002, Chapter 3). The story as it relates to self-control is a bit more complicated since the relationship between moral responsibility and self-control is itself, on Arpaly's view, predicated on preconditions these ideas share. But, moral responsibility and her conception of self-control

might reasonably be expected to vary together as functions of these preconditions, and autonomy might be expected to be affected correspondingly because of its relationship with self-control.

Thus, based on the meanings of autonomy that Arpaly identified as being relevant to moral responsibility, many of our judgements about these two ideas could be expected to lead in similar directions, which could have contributed to creating the impression that one depended on the other, or that they were the same thing. The benefit the multi-level of analysis is that it allows us to examine in detail the places where our intuitions about autonomy and responsibility might lead us in the same direction, while preserving the distinctness of the concepts. But, if there was some other precondition that was relevant to both concepts but played a different role in each[51], then we could also see that an overlap between two concepts does not necessarily lead to similar judgements about them.

I would also contend that similar kinds of attention to the second and third levels of analysis enhances our ability to make judgements about the plausibility of a wide range of claims. This is especially true of proposals pitched at the first or meta-level of analysis. An example of a meta-level trait is integrity. Some philosophers have claimed that autonomous people will have integrity[52]; they will stick to their principles in the face of opposition or adversity. The multi-level approach I have suggested provides resources for evaluating such claims because it forces us to be more precise about autonomy. So long as we can be sufficiently precise about what is

---

[51] This would also be true if it was discovered that the roles of the meanings of autonomy endorsed by Arpaly was not so similar in each concept as I proposed. This might be true to some extent in the present case in the sense that moral worth has two poles—praise and blame. If we shift the focus to blameworthiness, then a deficit in reason responsiveness would have a negative impact on our assessment of someone's autonomy, but increase their blameworthiness.

[52] A claim that Arpaly would possibly also find too vague to parse.

meant by integrity, an analysis informed by a multi-level approach can contribute to a precise assessment of claims of this type.

For example, different accounts of autonomy might suggest that we should expect different degrees of integrity from autonomous people. One account could propose that autonomous people will demonstrate absolute integrity, while another might suggest that autonomous people will merely demonstrate a greater degree of integrity than non-autonomous people (or perhaps more appropriately people with less autonomy). We can then use our common-sense understanding of psychology and our second and third levels of analysis to make judgements about which proposal is most reasonable. Cognisant of human frailty, and conscious of the ways in which social pressures can undermine the effectiveness of our means for achieving autonomy, we would be inclined to think that by adequately exercising autonomy in particular circumstances people do not simply become impervious to social pressures. For instance, someone who has been gaslighted will come to lack self-trust, and may therefore at times not be self-realising. In lacking self-trust she may fail to act on her principles because her conviction that they are right has been destabilised. While being gaslighted does suggest a deficit in autonomy, at least in some contexts, having a deficit in just one part of autonomy does not make someone entirely un-autonomous. Thus, a proposal that people who exercise comparatively higher levels of autonomy will demonstrate a greater degree of integrity, and will be better able to resist related threats to autonomy to some extent, is more appropriate than one that claims that people who exercise autonomy will possess unfailing integrity. [53]

---

[53] This is both because it is more realistic, which is my point here, and because unfailing integrity seems to call for an heroic conception of autonomy and is therefore bound to be unattainable and not a good candidate for inclusion in an adequate account of autonomy.

In this chapter, I have outlined my multi-level model for conceptualizing and analysing autonomy and articulated how it promises to improve our understanding of autonomy. I have shown that analysing autonomy as a multi-level concept (and phenomenon) allows us to rethink how problems relating to autonomy are addressed and provides the basis for more satisfactory explanations and judgements. In the next chapter, I will turn to problems manifesting themselves at the micro-level of analysis. Specifically, I will address whether, from a common-sense psychological perspective, we should worry that nearly everyone has significantly limited autonomy, given the number of capacities and other factors that are required for the exercise of autonomy. I will argue that this concern is less pressing than we might imagine due to special features of the micro-level of analysis, which shape how meso-level functioning can be attained.

Chapter 4: Extensiveness is Compatible with Practicability

Whereas the previous chapter contended with the problem of whether the *extensiveness* of autonomy leads to incoherence, this chapter will address the remaining potential concern about many-featured conceptions of autonomy that came into focus at the end of Chapter 2. Namely, the worry that understanding autonomy as ultimately grounded in a considerable number of skills, capacities, and attitudes makes achieving autonomy too complicated to be realistic for ordinary people. A potential implication of the sheer diversity of the factors that are relevant at the micro-level analysis of autonomy is that the practicability and therefore plausibility of autonomy is undermined. If this were true, then conceptions of autonomy that make use of them will not satisfy this important measure of theoretical adequacy. But, I will contend that this problem can be diffused by recognizing certain special features of how elements of the micro-level can combine to produce meso-level functions. In developing an account of these features— degree, automaticity, and reinforcement and substitution— and spelling out their implications for the exercise of autonomy, I will show that there is good reason to believe that while a complex story needs to be told in order to give a full account of what autonomy is and how it is exercised this does not mean autonomy is overly difficult to achieve. The three features I have enumerated, I will argue, are central to establishing that the demandingness of autonomy falls within the bounds of a common-sense understanding of human psychological limitations.

I will begin this chapter by synthesizing my earlier discussions about the capacities and attitudes required for autonomy from Chapter 2 with my discussion of the micro-level of analysis from the previous chapter to develop a more in-depth portrait of autonomy at the micro-level. In the next section, I fully articulate the potential problem with autonomy that arises from the way I have portrayed the nature of the micro-level of analysis. As I have already suggested, the concern is

that a full enumeration of the elements under consideration at the micro-level would suggest that autonomous people would need to possess quite a wide range of skills, abilities, and attitudes to exercise autonomy. This is consistent with worries about autonomy raised by Laura Davy, and mirrors certain difficulties faced by Martha Nussbaum's Capacities Approach. Theoretical resources from these authors, along with examples of how an insufficiently developed account of the micro-level analysis would reveal compromised autonomy in people with relatively common characteristics, will help to motivate an appreciation of the force of the concern about the implications of the *extensiveness* of the concept of autonomy for its practicability.

In the third section, I begin my positive case in defense of the practicability of autonomy. My defense relies on supplementing my account of the micro level with three additional features: degree, automaticity, reinforcement and substitution. Together they will act as the basis on which I alleviate concerns about the practicability of autonomy. In this chapter, I will give an account of the first two features[54]. Although these will bear strong resemblances to strategies employed to resolve common-sense problems with autonomy discussed in the first section of Chapter 2, there are some important differences. First, the elements of autonomy at issue are now located within the multi-level framework. Second, the resolution strategies that are being added in this chapter as new features of the multi-level approach are being employed at this stage to speak to a much broader number of aspects of autonomy. Finally, these features will not just demonstrate

---

[54] A discussion of the final feature of the micro-level that reduces the demandingness of autonomy, *reinforcement and substitution*, will be reserved until Chapter 5. This in my view is the least well recognized process for facilitating the exercise of autonomy. On my account, deficits in micro-level skills can be overcome by reinforcing them or substituting for them. This would be accomplished by leveraging other skills and attitudes or, importantly, extra-personal resources.

the consistency between autonomy and human flaws, but will be presented as means for enhancing autonomy.

Degree has been part of our discussion since the outset. But, at this point I will pause to give a more developed explanation of how an adequate conception of autonomy does not require that people are perfect with regard to the skills and attitudes that are considered at the micro-level. To be autonomous it may only be necessary to achieve some threshold level in any of them.

Automaticity was treated to some extent in chapter 2 in the context of responding to the criticism that autonomy could require more introspective ability than we would generally think people capable of. After briefly rearticulating the account of automaticity provided in that section, I will provide some examples of practices and processes that are consistent with our common-sense understanding of psychology and in which automatic responses are continuous with or facilitate autonomy.

Throughout the third section of this chapter, where I discuss degree and automaticity, I will rely on two principal examples of people with potentially reduced autonomy to illustrate each of the additional features of the micro-level. These examples will be central to the second half of this chapter, and will continue to feature in Chapter 5.

A Reprise of Autonomy Skills and Attitudes at the Micro-Level of Analysis

On my three-level model for examining and understanding autonomy advanced in the previous chapter, the micro-level of analysis represents the most granular level at which autonomy can be analysed. Although the relationship between this level of analysis and the evolution of what I have called the basic procedural conception of the capacities required for autonomy described in the section *A New Common-sense Concern* in Chapter 2 has been hinted at in my depiction of the micro-level in the previous chapter, it is worth identifying this relationship explicitly. It is my

contention that the suggestions of basic capacities or abilities involved in autonomy outlined in positive accounts of autonomy as advanced by John Christman, Diana Meyers, Catriona Mackenzie, Trudy Govier and many others, as well as accounts of how autonomy is undermined or put at risk, are for the most part considerations that would be taken up at the micro-level of analysis. As such, we can expect that the skills and attitudes required for autonomy on the micro-level of the three-level model are much the same as those that have been advanced as being involved in establishing authenticity, i.e. the important relationship between people and their self-defining attitudes, and supporting competency, i.e. the process of living according to our self-defining attitudes. The attitudes that would be relevant considerations at the micro-level of analysis include at least those that are normally required for autonomy. Strong cases have been made that self-esteem, self-trust, and self-respect fall into this category.

So, there is significant overlap between the content of the section of Chapter 2 named above and what we might expect to discuss at the micro-level of analysis. However, further deliberation and debate would be required to say decisively which attitudes and skills are part of a definitive or correct conception of autonomy. I will tend to discuss and employ those that I think are most promising, but will not advocate for a definitive or comprehensive set of skills and attitudes, largely because I'm sure I have missed some possibilities that may indeed find a place in a correct conception of autonomy.[55] I have provided at least enough resources to convey the general sense of the kinds of factors that could be relevant at the micro-level of analysis.

---

[55] And, indeed, it may be the case that in some ideal conception of autonomy some of the items I think should be reasonably included will not survive.

To reiterate from Chapter 2, at a minimum the following have been advanced[56] in contemporary

theorizing as having a role in the exercise of autonomy:

- Rationality (Dworkin 1976, 27, Christman 1991, 16)	- Comprehension (Mackenzie 2014, 19)
- Judgement (Dworkin 1976, 25)	- Deliberation (Meyers 1989, 55)
- (Critical) Reflection (Dworkin 1976, 26, Meyers 1989, 81)	- Self-respect (Mackenzie 2008, 525, Wolf 1989, 144-45)
- Introspection (Meyers 1989, 47, 79)	- Retrospection (Meyers 1989, 53)
- Emotional attunement to self and others (Meyers 2005, 46, Mackenzie 2014, 34)	- Sanity (Meyers 1989, 62)
- Imagination (Meyers 1989, 78, Foddy and Savulescu 2010b, 35)	- Memory (Meyers 1989, 79, Meyers 2000, 166)
- Volition (Meyers 1989, 79, Mackenzie 2014, 18)	- Resistance to external pressures (Meyers 1989, 83)
- Determination (Meyers 1989, 83)	- Curiosity (Meyers 1989, 87)
- Communicative abilities (Meyers 1989, 87, Meyers 2000, 166)	- Reliable perceptual abilities (Killmister 2013, 112)
- Interpersonal skills (Meyers 2005, 38, Mackenzie 2014, 34)	- Embodied practical intelligence (Meyers 2005, 40)
- Decisiveness (Killmister 2013, 113)	- Self-control (Killmister 2013, 114)
- Motivational competence (Christman 2014, 373)	- Self-trust (Killmister 2013, 115, Mackenzie 2008, 525)
- Self-awareness (Christman 1991, 17)	- Self-worth (Govier 1993, 110)
- Disposition to answer for one's self (Westlund 2003, Benson 2005)	- Dialogical skills (Mackenzie 2014, 34)

This list, which I have admitted is likely not exhaustive is already quite extensive. The

relationship between the micro-level of analysis and the meso-level of analysis suggests that the

---

[56] These are the same citations as offered in Chapter 2, which are only examples of where these capacities appear in philosophical treatments of autonomy.

simultaneous activity of many of them is what is required to produce the meso-level functions. I have already given a description of what I view the relationship between the micro and meso-level of autonomy to be like in chapter 3, but I will reiterate it here briefly. We can expect that there is a cluster of skills and attitudes that make each meso-level function possible. In the previous chapter, I employed the example of self-definition. By reasoning through the kinds of activities that would be involved in self-definition-- such as considering the appeal of particular self-defining attitudes and making judgements about whether to adopt or reject them, and knowing about the world—we can construct a list of means that would be required if someone is to define themselves. I proposed that self-definition would require skills involved in self-knowledge, which with reference to our longer list above we might conclude involves: introspection, retrospection, rationality, memory, curiosity, self-awareness and dialogical skills. Furthermore, insofar as self-definition involves deciding whether to adopt particular self-defining attitudes, it will also require list items like: judgement, critical reflection, perceptual skills (since at least some pragmatic considerations must be relevant), self-worth, self-trust, imagination, deliberation, and emotional attunement to one's self.

Catriona Mackenzie's discussion of the role of imagination in self-definition illustrates how a deficiency in or absence of one or more of the skills required for self-definition could undermine this function significantly. She argues that imagination plays important roles in how people construct their identities, i.e. define themselves (Mackenzie 2000, 133). She claims that the emotional valence of what we imagine can initiate introspection, and through its affective force inform judgement (Mackenzie 2000, 136). She also claims that this emotional dimension of imagination—especially in the form of "counterfactual speculation [or] future directed fantasy"— can also help us know ourselves by helping confirm or reveal discrepancies in our

judgments about what is important (Mackenzie 2000, 137). Our imaginations also allow us to consider and explore new possibilities, giving us a sense of who we might be if we were to give more or less prevalence to certain of our ideals or characteristics relative to others. This process can also be used to anticipate whether we might endorse whole new ways of life (Mackenzie 2000, 138). Her account gives us reason to believe that imagination is an important complement to the other means (i.e. skills, attitudes, possibly other factors) involved in self-knowledge such as self-awareness and introspection, as well as to judgement, deliberation, and emotional attunement.

Mackenzie, however, is also attentive to the ways in which our powers of imagination can be undermined or impoverished by lack of information and oppressive social conditions (Mackenzie 2000, 138, 143). She is particularly concerned about the extent to which the effects of the latter on imagination "impair [people's] capacities for self-definition" (Mackenzie 2000, 143). It's clear how this would be the case since imagination at the very least enhances, if not provides the basis for the self-understanding and decision making that is involved in this function.

I would contend that the effects of a lack of imagination could be even more wide-ranging. This is because a deficit in imagination could equally undermine other meso-level functions that are the basis for autonomy. My thought is that just like self-definition, other functions will be underpinned by their own sets of micro-level means. This leaves open the possibility that a set of means associated with one function may share members with the set associated with a different function. For example, I would suggest that imagination plays a role in practical reasoning and therefore in functions such as external self-realization/self-determination. The affected processes would not be those that relate to the largely mental task of establishing the authenticity of our self-defining attitudes, but rather those involved in living in accordance with them. Although

there may be different conceptions of what it means to live according to our self-defining attitudes[57], only some potential courses of action will count. In order to develop a plan for satisfying our desires or achieving our goals, we at least sometimes have to be able to imagine the different possibilities that would count as realizing those goals, the various courses of action that could lead to them, and their likely outcomes.

It is my contention that we can generalize the case of imagination to understand the ways that the skills, attitudes, and other means at the micro-level influence functioning at the meso-level. Indeed, the effects of diminishments or impairments in means other than imagination may be even greater. Think about rationality, reasonableness, or judgement. These abilities are essential to both authenticity and to acting in accordance with our values. Someone who is limited with respect to them will have a distinctly difficult time exercising autonomy because of their shared importance to a great many meso-level functions. As such, the absence of a particular skill or attitude at the micro-level has the potential to affect several functions at once, meaning that such an absence or impairment— even of a single skill or attitude— can represent a significant blow to someone's ability to exercise autonomy.

The picture of the relationship between the micro-level and meso-level I have outlined involves an interconnected web of functions on one level, and skills, attitudes and other means on another. Different functions rely on distinct, but potentially overlapping sets of means. Different clusters of means support different functions, but because of their overlapping members, the operation of some functions is not entirely separate from the operations of others. Our ability to achieve one

---

[57] For example, it can be debated whether living according to your authentic beliefs and values requires causing certain effects, or whether we need to be concerned about effects at all. We might accept it is at least sometimes possible to live according to our values based purely on the nature of an act or the intention behind it with no regard to the effects or outcomes.

may depend on factors similar to those that allow us to achieve others. In the next section, I will consider a worry that may emerge from this depiction of the way autonomy is exercised. The problem is that not only do autonomous people have to be capable of a range of autonomy functions, but these functions themselves rely on the simultaneous coordination of many other factors. So, exercising autonomy is a very *complex* process that involves an *extensive* number of constituents even if we think that the functions more closely related to establishing authenticity do not have to be at work at the same time as the functions for acting on authentic self-defining attitudes. And, deficits and impairments in the most basic precursors to autonomy at the micro-level can have far ranging effects on our overall ability to exercise autonomy. Many skills and attitudes need to be recruited into the meso-level functions, and each of them is an opportunity for a misfiring or failure that has potentially widespread effects through the entire system. These observations might lead us to the conclusion that the theoretical work that has gone into making autonomy reflect human realities more accurately has simultaneously made it less plausible that autonomy is widely enjoyed.

The Problem of Many Means

In this section I will develop the worry about the *extensiveness* and *complexity* of autonomy and its implications for practicability more fully. As stated above, the general objection is that there are *so many* different factors at play in the exercise of autonomy and so many ways that it can go wrong that it is hard to believe that most of us, and especially those of us with physical or psychological conditions that involve impairments in the skills and attitudes that belong to the micro-level, should qualify as autonomous. Or, more generally, most of us probably experience some form of impairment with regard to the elements of the micro-level whether or not these have been recognized by medical or other professionals as belonging to an identifiable pathology or disorder. It would be no surprise to find that people who function within the normal

128

parameters of human life are very often somewhat impaired when it comes to the micro-level skills and attitudes. Impairments may range from minor cases of reduced self-trust or self-esteem to more profound challenges relating to major deficits in motivational effectiveness or judgement, or greater or lesser limitations with respect to several micro-level elements. So, the question becomes what does it take to meet the criteria for autonomy, and does that agree with what we think it means for autonomy to not only be possible but plausible?

Martha Nussbaum faced similar questions in regards to her capabilities approach to ethics. The development of this approach begins, for her, with an investigation of what a good life for humans requires by thinking about what is held in common by all people through human experience (Nussbaum 1995, 62, 69-70). She arrived at ten capabilities she thought were central to human life and essential to a dignified human existence (Nussbaum 2011, 32-4). Given that her overall view is that the achievement or attainment of human capabilities structure political obligations (Nussbaum 1995, 82), what she argues it means to be a human being has both conceptual and normative implications. "[G]etting the concept right" is important not only because it distinguishes between people in a way that hopefully tracks a property we are interested in, but also because ethical and political status attaches to being found to qualify as a being of the appropriate type (Nussbaum 1995, 71-2). So, the stakes are high not only because of the desire to have a meaningful theory, but also because of the implications for social relationships. The process of arriving at a plausible conception of autonomy faces similar challenges and stakes. Our concern with theoretical adequacy in the form of practicability is significant both conceptually *and* politically. Getting our account of autonomy right is important not just for its own sake but also because it *matters* in the ways that it promises (or threatens) to shape the courses of peoples' lives.

Despite the evolution of contemporary thinking about autonomy to embrace proposals that it is at least partly relational and that it relies significantly on many mental capacities in addition to reason, Laura Davy has recently charged that even feminist conceptions of autonomy remain too individualistic and potentially too cognitively demanding to be acceptable models of autonomy for people with certain kinds of cognitive disability (Davy 2015, 139-40). She claims this is in part because of the many skills that have been proposed to underlie the exercise of autonomy, as well as the retention of the expectation that autonomy is ultimately an individual achievement (Davy 2015, 139-40).[58] Davy's concern is that even taking into account the contributions of scholars from feminist and disability studies, going accounts of autonomy and its role in social life fail to "accommodat[e] human diversity" (Davy 2015, 132). I find Davy's worry here compelling, but for perhaps slightly different reasons than the ones that preoccupy her. For Davy, the project of expanding the applicability of autonomy to people with a wide range of attributes is motivated by a concern for social justice. She wants to resist the segregation and stigmatization of people with disabilities by making autonomy more accessible or inclusive (Davy 2015, 132).

My concern is that the influence of her underlying motivations on her arguments begs the question, since she begins with the supposition that people with intellectual disabilities are entitled to the social standing associated with autonomy, and proposes modifications to our understanding of autonomy that support this conclusion. Her view might therefore reasonably be cast as the kind of view to which Mark Piper (who we encountered in Chapter 1) objects when he claims that developing one's account of autonomy with the aim of achieving certain "desirable

---

[58] The other aspect of the view Davy advances is that people with intellectual disabilities and their carers are bound together in negotiating "a social matrix of needs" in which the autonomy of the former may easily come to threaten the autonomy of the latter. This is a supremely interesting point, but is beyond the scope of the present discussion, I hope to engage it in future work.

normative implications" such as maximizing the scope of our duties involving autonomy-based respect "put[s] the cart before the horse" (Piper 2016, 286-7). He persuasively argues that considerations of what might be called 'solidarity' are not relevant to establishing who does or does not have the property of being autonomous. My reservations about Davies' view are less pronounced than Piper's. In Chapter 1, I defended the idea that a satisfactory account autonomy would concern itself with the practicability of exercising autonomy at least partly for socio-political reasons. But, I also proposed that the nature of social practices could require non-trivial levels of functioning to be met in order for someone to qualify as autonomous in that context.

So, although Davy is entirely correct in indicating that the social stakes of our inquiry into the nature of autonomy are very high, developing a reasonable account of autonomy does not rule out the possibility that some autonomy thresholds may be beyond the grasp of some humans. Attempts to claim otherwise, i.e. to guarantee everyone qualifies as autonomous results in some strange propositions. For instance, Nussbaum proposes (and this proposal is repeated by Davy) that part of making possible the capability of political life in people with cognitive disabilities would be for their political rights (to vote for example) to be exercised by a surrogate "even if the person is unable to make a choice" (Nussbaum 2011, 24, Davy 2015, 135). Though Davy sees herself as providing a justification for such practices from the perspective of autonomy, I would question whether this qualifies as an exercise of autonomy by the agent on whose behalf these rights are exercised.[59]

That is, while there may, as I will argue, be many ways in which a person can come to make a choice and therefore partly satisfy the functions that constitute the exercise of autonomy, to meet

---

[59] Indeed, we might find Nussbaum's articulation of this point less dubious precisely because she does not base her claim on autonomy, but rather on citizenship and other factors (see also Nussbaum 2005).

an autonomy threshold they must exercise all of the relevant functions, which does not appear to be the case in extreme examples. While Davy is right to criticize conceptions of autonomy that put too much emphasis on independence, my own research with people with (albeit mostly physical) disabilities suggests that there is something important in at least being involved in carrying out one's life activities for these activities to be meaningfully one's own. In other work, I report on interviews I have conducted with people who access home and community support services. During these exchanges a recurrent theme was how important it was to them to have the opportunity and latitude to participate in the activities with which they were being assisted to the extent they were able. This includes people who were getting help bathing, cleaning, shopping, and other basic activities of daily life. To me it seems like the case of voting on behalf of someone who is unable to even choose how to vote does not represent an instance of the exercise of autonomy since there appears to be no meaningful participation—in the sense described in Chapter 1-- of the individual on whose behalf the vote is being cast.[60]

So, I do not think that the model of autonomy I am developing can claim to apply to all humans despite Davy's proposal that autonomy should not be seen as excluding any person from entry (Davy 2015, 143). But, in explaining how autonomy, despite the large number of factors involved, is still plausible for most people, I suspect that I will provide the grounds for maintaining that many more people may be capable of autonomy than is generally recognized

---

[60] Now, allowing surrogate voting might be justified under different understandings of what voting is all about. For example, if we understand voting as being primarily a matter of having one's interests represented, then surrogate voting may be more justified since a person can have interests even if she doesn't have the capacity to choose. But, then the argument in favour of this practice should not be made on the basis of autonomy. Thanks to R. Myers for this point.

provided they have access to the right supports and opportunities. I imagine this would include people with at least certain degrees of intellectual disability and other cognitive limitations[61].

Given that the standard of practicability I am trying to meet in this account is oriented toward explicating the intuition that people are for the most part autonomous[62], Davy and I share the concern that procedural[63] (and other) conditions outlined in recent thinking about autonomy not only unjustly exclude people with certain disabilities, but also "fall short of modeling how *the vast majority of the population* exercises autonomy"[64] (Davy 2015, 139). With questions about justice in the backs of our minds, we cannot help but wonder whether the number of people who are able to deploy all the skills, capacities, and attitudes required for autonomy is significantly more restricted than our intuitions about autonomy would suggest. To be sure, any account of autonomy that denies that any humans whatsoever are autonomous will ultimately have to face challenges about the boundaries of the concept advanced by people like Davy. But, my present concern is that without further amendments to the understanding of the exercise of autonomy I began to advance in chapter 3, and which was more fully elaborated in the first section of this chapter, we might reasonably think that many autonomy thresholds exceed even what we would count as well within the range of normal human functioning and living a normal life. So, while I will not attempt to address the more remote boundary case of severe intellectual disability in this chapter and the next, I will seek to understand how people whose autonomy (according to the view above) could be perceived as impaired, may nonetheless still often be able to exercise adequate levels of autonomy.

---

[61] Hopefully, this will help resist Licia Carlson's charge that cognitive ableism is ingrained in philosophy.
[62] As suggested in chapter 1 in the section entitled *The Importance of Empirical Considerations for Autonomy*.
[63] Davy's target in this context is Diana Meyers's paper from 2000 in which she lists seven skill sets that autonomous people must possess; these also inform the list presented earlier in this chapter.
[64] Emphasis mine.

To help illustrate the problem I am pointing to and the amendments that will help resolve it, I offer two relatively quotidian examples of cases in which we would not want to rule out the possibility of autonomy. This reluctance would persist, in my estimation, even though the people in the cases lack at least one skill, capacity, or attitude likely to be important in a micro-level analysis of autonomy. The cases are those of a frail elderly person living alone and an adult with ADHD.

Case 1

Dolores is one of the more than a million older adults who live alone in their homes in Canada.[65] *She has health conditions that affect her respiration and mobility, and so she uses an oxygen tank to help her breathe and a walker to move around. She lives in an apartment in a small town. Because of her fragile health condition she doesn't trust herself to bathe unassisted and rarely goes out, especially in the winter, for fear of falling and injuring herself. This lack of confidence in her abilities has started to shape not only how she acts, but the reasons why she acts. She has started to re-evaluate what's most important to her in light of her assessment of her abilities. For example, she has given up on trying to take walks outside for exercise, and instead attempts to exercise by walking the halls of her apartment building even though she finds such walks less interesting and satisfying. And, she finds herself making sure to pursue her plans in ways that accommodate what she perceives as her limitations. For instance, despite only having a small apartment she invites family members who live far away to visit her, since her oxygen tank only allows her to make short day-trips.*

---

[65] See Statistics Canada 2017.

To what extent is Dolores autonomous? Our preliminary assessment of her situation on the micro-level of analysis would note her lack of self-trust as potentially undermining her autonomy. It may be worth noting that her lack of self-trust itself stems from certain kind of loss of physical self-control, a decrease in her ability to independently transform her decisions and motivations into action due to a loss of bodily strength and coordination. The paragraph above suggests quite strongly that these deficits are having a negative impact on her ability to exercise in a way that is consistent with her values. The further references to the changes that she is undergoing as they relate to her judgements about what is important in life and how she goes about living according to what she finds important also raise preliminary questions about whether she is developing adaptive preferences. Her lack of self-trust appears therefore to be having effects on both the meso-level functions involved in establishing authenticity—self-definition for instance-- and the meso-level functions involved in acting authentically, such as self-realization. So, we might conclude that her autonomy is being undermined by her lack of self-trust or self-confidence since, for example, it seems that if she had her way she would continue walking outside, like she used to.

Case 2

While he generally understood what he was being taught in school, Egon has always had difficulty turning that knowledge into action. With tutoring and close parental monitoring he was able to complete high school and continue on to college. However, he continued to struggle in college because he had difficulty sitting through and maintaining attention in lectures, had difficulty establishing a full understanding of and prioritizing his assignments, and was easily distracted when studying or completing coursework. Difficulties of this nature are also manifesting themselves now that he has entered the workforce, and he constantly worries he is

135

on the brink of being fired for poor performance and difficulty managing his daily

responsibilities. He feels he is also struggling in his home life, where he experiences regular

conflict with his domestic partner when he forgets to do his share of the chores because he is

absorbed in something else, or fails to stick to their household budget because of a tendency to

buy things impulsively. As a result, Egon is not achieving the success he would like in his career,

and is not the partner he would like to be in his relationship. Upon seeking psychological

assistance he is diagnosed with ADHD.[66]

The details of Egon's case raise doubts about his ability to exercise autonomy, and therefore be autonomous. Underlying the behaviours he has—his ADHD symptoms—is a developmental delay that first manifested in his childhood and which continues to affect his ability to function at a developmentally 'normal' level. Some researchers trace these symptoms back not merely to inattention and hyperactivity, but rather to a more fundamental deficit in executive functioning that has a pervasive effect on individuals' ability to self-regulate (Barkley 2000, 12). So, although Egon seems to have personal values and goals that he is trying to pursue, such as being a good partner, furthering his education, and performing well at work, his inability to exercise the relevant degree of control serves as an impediment. Our concern is not about the authenticity of Egon's values[67], but instead whether he has the micro-level means necessary to be self-realizing or self-determining to a satisfactory degree in at least everyday circumstances.

---

[66] This profile was informed by Barkley, R.A., Murphy, K.R., and Fischer, M. *ADHD in Adults: What the Science Says.* 2008. The Guilford Press: New York.

[67] At least not with the case as described. It is possible that further plausible details could be added to convey how his attitudes and aspirations have been shaped by the way ADHD has frustrated his attempts to exercise his will, potentially making his preferences and values adaptive.

If Dolores and Egon, who represent not terribly unusual cases[68], were found on a multi-level analysis to have such severely limited autonomy that they rarely met everyday autonomy thresholds then we may be inclined to worry that my multi-factor model of autonomy is overly exclusive since it makes autonomy overly extensive and possibly overly complex. There are many degrees of reduced ability between their cases and the usual excluded cases such as people in comas, babies, and people who are insane. If Egon and Dolores are deemed to be generally unable to exercise autonomy, then potentially people with many other relatively common conditions will be too. The purpose of the remainder of this chapter is to argue that, with suitable development of the multi-level model of autonomy I have advanced so far, we needn't think that Dolores's physical limitations necessarily lead her to fail to qualify as autonomous. Nor, I will argue, would we be right to believe that Egon's ADHD is a necessary limitation on his ability to exercise autonomy. In articulating how Dolores's and Egon's cases are consistent with autonomy, I will be responding to the charge that, even if it is coherent, the model of autonomy I am advancing is too demanding for average people to reasonably be expected to meet. Most people, like Dolores and Egon, are probably not fully masters of the micro-level autonomy skills and attitudes that are the basis for the meso-level functions. As such, most of us are likely lacking when it comes to the meso-level functions that together make up autonomy. In what follows, I will articulate the kinds of conditions under which both of these individuals, as well as individuals with lesser and perhaps also greater limitations, could be able to exercise autonomy. The final model should be one in which our confidence that people are largely autonomous is restored from a common-sense psychological perspective.

---

[68] Older adults living at home alone and adults with ADHD each appear to make up about 4-5% of the population (see Barkley, Murphy, and Fischer 2008, p. 23 and Statistics Canada 2017).

Solutions to the Problem of Many Means

There are, as I see it, two major avenues for helping defend the plausibility of autonomy from the concern that it is overly demanding. Both, I think point to the need to more fully articulate important features of the multi-level model for the purpose of understanding what autonomy involves, particularly from the perspective of the micro-level. The first is to temper expectations about what level of mastery people are expected to exercise with respect to the micro-level means. The second is to entertain the possibility that the meso-level functions might be achieved by multiple combinations of means, revealing flexibility in, and further resources for, the exercise of autonomy. In elaborating the model along these two lines I will show that micro-level capacities and attitudes, and ultimately autonomy itself, to be a matter of degree. I will also support the idea that people can rely on automaticity, reinforcement, and substitution to facilitate the exercise of autonomy. These additional features of the model of autonomy constitute the provisions that make it a convincing non-ideal theory of autonomy, which is necessary for plausibility because I am attempting to accommodate the foibles and imperfections of human agents in my theorizing (Mackenzie 2014, 24). The cases of Dolores and Egon will be employed to show how these further features of the micro-level pave the way toward showing how even people who experience limitations associated with micro-level attitudes and skills can exercise autonomy.

In this chapter I will discuss the first avenue, that of tempering expectations surrounding autonomy, and the first part of the second avenue which considers how the exercise of autonomy can be facilitated through automaticity. I will reserve discussion of reinforcement and substitution for the final chapter of the dissertation, Chapter 5. I have divided things in this way for two reasons. First, the roles of degree and automaticity in theorizing about autonomy have already been relatively well established in the work of other theorists. So, the measures for

reasserting the practicability of autonomy I will address in this chapter already play important roles in theorizing about autonomy, as demonstrated by the fact that they have already appeared in earlier sections of this dissertation. By contrast, the roles of reinforcement and substitution are not well recognized.[69] Giving them essentially their own chapter highlights that they are one of the more novel aspects of my contribution. Second, understanding that autonomy is a matter of degree and that it can be made easier to exercise through automatic processes does not do much to address the question of how autonomy can be exercised when someone faces relatively significant limitations in their personal psychological abilities. The amendments to my account of the way the micro-level works discussed in this chapter still depend on significant personal psychological abilities. The amendments that are of central interest in the next chapter will address how autonomy might be exercised when one's personal psychological abilities cannot be relied on.

i.   *A Matter of Degree*

The first way in which the multi-level model of autonomy requires further elaboration highlights the extent to which the micro-level functions, and therefore autonomy itself, are a matter of degree. That I should include this as a feature of my model should be no surprise given the course of the discussion in Chapter 2. In Chapter 2 we saw that many of the common-sense concerns about autonomy relate to whether autonomy would be possible or plausible if it demanded perfect rationality, self-knowledge, or self-sufficiency. In all cases we found autonomy theorists denying that autonomous people were expected to be perfect in these regards and, among other things, proposing that if these were necessary for autonomy at all, they were

---

[69] Though I acknowledge that they are sometimes alluded to or mentioned in passing. Take for example, McLeod and Sherwin's claim that the self-trust of women with addictions can be reinforced relationally in their relationships with healthcare workers (McLeod and Sherwin 2000, 265).

only necessary to some limited extent. Then near the end of chapter 2, I articulated the ways in which the basic proceduralist picture of autonomy had been elaborated upon to give a more fully developed account of what kinds of activities and functions authentically selecting our self-defining attitudes and living according to them would involve. I noted that a potential concern facing a conception of autonomy that includes these additional skills and attitudes would be that if it expected perfection on all these fronts, it would be concerningly implausible since common-sense would tell us that everyone can be expected to have flaws or lapses in one or more of these areas. At that point I indicated that many theorists responsible for contributing to the increased complexity of autonomy also accepted that autonomy is a matter of degree. In this section I will examine this claim more closely, contending that autonomy is correctly viewed as a matter of degree at least in part because this reflects the common-sense expectation that an individual may only imperfectly carry out the functions that make up the meso-level, which could potentially be traced to further limitations at the micro-level of analysis.

One source of the concern that autonomy may require that our micro-level attitudes and skills exceed what we might reasonably expect of people arises in imagining what might be necessary in order for our meso-level functions to achieve what they should. Catriona Mackenzie raises this concern with respect to self-authorization. For her, self-authorization involves "regarding oneself as having the normative authority to be self-determining and self-governing", that is to have a privileged stance towards the adoption of one's priorities and the project of carrying them out in the world (Mackenzie 2014, 36). Among other things, Mackenzie takes self-authorization to involve holding self-evaluative attitudes such as self-trust and self-respect, potentially even in a context where social reinforcement of these attitudes is not forthcoming. She acknowledges that an understanding of self-authorization that takes this position will be seen by some as overly

demanding, requiring levels of self-assuredness that are unrealistic and that are not likely to be achieved by many people (Mackenzie 2014, 39). Part of her solution to this issue is to propose that conditions for self-authorization are matters of degree. So, her argument suggests that the functions underlying autonomy are not all-or-nothing in nature, but rather that lesser degrees of achievement can be adequate for autonomy.

The general picture that Mackenzie paints in this argument can, in my view, be aptly extended to the other meso-level functions. The various functions at the meso-level are postulated as playing particular roles in either the authentic adoption of values or living in accordance with them. People may perform the meso-functions to different extents or degrees, but for each function, the level required to achieve autonomy falls below the level of perfection. Accepting that less than perfect skills and attitudes at the micro-level, and less than perfect functioning at the meso-level, can still be adequate for autonomy, suggests that the exercise of autonomy itself might be a matter of degree (Beauchamp 2005, 316). Those with the most developed meso-level functions would presumably be more proficient in the exercise of autonomy than those functioning at lower levels. This possibility raises important questions and problems relating to setting a threshold for who counts as autonomous, what degree of competence in the exercise of autonomy is sufficient. There are two central questions about a prospective autonomy threshold, so far as I can see. First, there is the question of what *kind* of threshold it is. And second, there is a question about the level of functioning at which the threshold should be set. The discussion in the rest of this section will address these questions and in so doing serve as a defense of the claim that autonomy is a matter of degree, which has so far gone without justification.

In regards to the type of threshold that we might employ when making judgements about autonomy when it is characterized as a matter of degree, there appear to be two options. One

141

option is that anyone beyond the threshold is to be regarded as autonomous. On this model of threshold, there is only one judgement to be made: does the person satisfy the requirements of the threshold or not. This kind of threshold is, as Joel Feinberg says, "all or nothing" in that it spells out necessary and sufficient conditions which people either meet, or fail to meet (Feinberg 1989, 28-9). And, those who meet these conditions qualify as having the property (in our case autonomy) specified are deemed to possess the property in an absolute rather than relative sense.

Feinberg argues that our thinking around political rights to self-government tends to employ this type of threshold. That is, on political views where being autonomous is the basis for a particular kind of status that involves the right not to be interfered with, as well as certain kinds of empowerments and responsibilities (Anderson 2014, 359), one does not have these rights to a greater or lesser degree depending on the extent to which one exceeds the threshold. Whatever capacity one has beyond the threshold is "simply unused surplus" (Feinberg 1989, 29). As I see it, there are two problems with viewing the threshold for autonomy in this way. First, as I have suggested above, it may actually be inappropriate to attempt to specify necessary and sufficient conditions for autonomy since autonomy is such a rich concept (Dworkin 1989, 54) rendering such an analysis inadequate to capture autonomy in its entirety (Mackenzie 2014, 18). So, establishing a hard threshold through the articulation of necessary and sufficient conditions might not capture all the phenomena of interest in making judgements about autonomy. Second, this all or nothing approach seems to make us blind to where people stand in relation to the threshold. This might not be such a problem on a view that holds that one's level of autonomy is innate or immutable, since there is nothing more to know about autonomy than one's level of it. A lack of interest in where people stand relative to a threshold for autonomy becomes a problem if we take autonomy to be something that can be continuously developed and improved. When

understanding autonomy in this way, knowing how someone fails to meet the threshold for autonomy, and in what way, will be of considerable interest from the perspective of guiding development toward the eventual attainment of the threshold level.

The second option is that the threshold is taken to represent a particular (hopefully significant but potentially conventional) position along a continuum of performance. On this model of understanding the threshold there is interest in the entire continuum and autonomy is not an all or nothing question. Instead, people's ability to exercise autonomy falls "on a scale from minimal to robust" (Piper 2016, 769). This, I think is the more appropriate way to conceive of thresholds related to autonomy.[70] One of my objections to the foregoing model of an autonomy threshold is that whether or not someone satisfies some necessary and sufficient conditions is not the only question of interest in making judgements about autonomy. Maintaining an interest in the broader picture of functioning relative to autonomy helps provide the grounds for celebrating and giving credit for the constrained performance of autonomy functions among those whose circumstances are not conducive to the fully fledged exercise of autonomy (Mackenzie 2014, 41), such as the deliberations, choices, and actions of women in conservative societies. These people are not non-autonomous as the all-or-nothing threshold might suggest, rather they are doing their utmost to exercise the degree of autonomy available to them in adverse circumstances (Narayan 2002).

Likewise, attentiveness to people's positions relative to the threshold even if they are beyond it allows for further judgements about whose autonomy is vulnerable or precarious (and why), or whose exercise of autonomy is more than adequate, but not what it could be. For example, in the

---

[70] I would argue that this is true even in the political context, although this argument would be quite involved and is beyond the scope of the present work. My hope is that the reader won't see me as changing the question as I transition from the first to the second kind of threshold.

former case we might note that Egon has enough motivational effectiveness to do what he needs to survive school while living with his parents, but he may lack sufficient motivational effectiveness to succeed in a less structured work environment or living independently. Loss of parental support or changes in life circumstances could move him from barely above the autonomy threshold to below it. Alternately, though we might see Dolores as a firecracker of an older woman who is naturally slowing down due to deteriorating bodily health, we might equally wonder whether, with some additional safety features in her home, she would be able to return to enjoying the activities that currently make her wary. Her lack of self-trust is influencing her outlook on life and shaping her actions and priorities, but not undermining her autonomy in a catastrophic way. Still, she may be able to exercise greater autonomy if the effects of her lack of self-trust could be alleviated in that she may be able to resist the reorientation of values precipitated by her decrease in self-trust (and self-control). Being sensitive to where someone falls on the spectrum is important because it is the way we recognize that (1) those who exercise only the bare minimum of autonomy may be at risk, and (2) those who exercise generally adequate levels of autonomy may still experience adverse and possibly morally problematic effects due to the limitations they face. To put the point very broadly, attending to the entire continuum or spectrum of levels of functioning related to autonomy provides us with increased insight into the nature of autonomy and the differences in how it is exercised by people above and below whatever threshold(s) we demarcate as salient.

The second major question when it comes to interpreting the exercise of autonomy as a matter of degree is, indeed, where to set the threshold for who counts as autonomous. In Chapter 2 we saw that demanding perfection or maximization of precursors to autonomy like rationality and introspection, which turned out to be micro-level skills, was rejected because it seemed to make

autonomy impossible. One way of interpreting this result is that the extreme standards implied a threshold that was too high to be plausibly achieved by anyone. I showed how autonomy theorists had worked to temper the extremes to develop more subtle interpretations of the relationship between these skills and the broader phenomenon of autonomy to which they contribute.

By contrast, however, Mark Piper is concerned that many leading theories of autonomy "contain problematically low thresholds for autonomy achievement" (Piper 2016, 772). Piper points to Marilyn Friedman's standard of minimal self-reflectiveness, Paul Benson's belief that the objective conditions of autonomy are "widely satisfied", Diana Meyers's requirement of retroactive resonance with one's actions, and John Christman's counterfactual endorsement of the process by which a self-defining attitudes are acquired as cases in point (Piper 2016, 772). Piper charges that on the whole these views of autonomy tend toward being too permissive when judging autonomy in that they do not seem to require the kind of control, the "focused act[s] of will" that he takes to be central to the notions of determining or governing one's self (Piper 2016, 773-4). However, as I discussed in Chapter 1, an understanding of autonomy that places it beyond human capability is not helpful for understanding human experience or the basis for certain kinds of political status. Likewise, a standard that serves, ultimately, as no standard at all—that classifies people as autonomous when they are not—is equally inappropriate. So, there is a real need to be precise about what counts as living according to our own values. My comments in Chapter 1 were intended as the beginnings of an answer to this question. Being adequately precise will help ensure that the right people are included and the right people are excluded. The question "excluded from what?", however, needs to be further addressed.

I also noted in Chapter 1 that autonomy is important both from a socio-political perspective as well as an ethical perspective. On the ethical side of things, much of the moral picture remains the same whether someone is autonomous or not. Principles other than autonomy provide the moral prohibitions against causing pain and suffering (Frey, 343). And, on a moral level rights to dignity and respect are possessed categorically, dependent neither on context nor on qualifying as autonomous (Mackenzie 2014, 40, Anderson 2014, 358). Furthermore, in understanding autonomy as a matter of degree, it seems like all people who are the least bit capable of the meso-level functions enjoy autonomy as a moral good to a greater or lesser degree in their lives. And, presumably, it is good when they can move up the scale toward higher levels of autonomy. As much as autonomy is part of a good life, it's not entirely clear how much bearing meeting a particular level of autonomy would have on one's moral worth. Instead, we might think that wherever someone falls on the autonomy spectrum it is good to contribute positively to their autonomy if possible, or at least not obtusely or maliciously get in the way of their exercise of autonomy. I would contend this is true even of people who are capable of only the least amount of meso-level functioning.

Still, we should not lose sight of the significant concerns about inclusion and exclusion in relation to autonomy that arise in the socio-political context where the "autonomy of individuals…impose normative constraints on the claims of social groups or collectives" (Mackenzie 2014, 21). For example, in exercising a sufficient degree of autonomy in a particular context one becomes entitled to have one's wishes taken seriously by others, especially in institutional contexts. By meeting the autonomy threshold your ideas about how to conduct your life are given certain social protection and priority. As I noted in Chapter 1, it is not clear that the

same standard should be applied in all of life's circumstances; to say otherwise is to risk some ridiculous implications.

In Chapter 1, I articulated what I take to be the bases on which autonomy thresholds should be set. These had to do with what it means for us to have meaningful ownership over our self-defining attitudes or the decisions that issue from them in particular contexts. My arguments in Chapter 1 were consistent with Mackenzie's claim that the importance of what level of functioning one attains, and the extent to which concern about deficits is warranted, will depend on these situational factors (Mackenzie 2014, 41). At that time, I illustrated how the stakes, rational and epistemic demands, and other factors might influence what degree of autonomy must be exercised to meaningfully participate or engage in particular social practices. In recognizing the varying demands for what it means for decisions to be truly our own I argued that we should follow those, like Mackenzie, who urge how important it is "to set different thresholds for meeting these conditions in different contexts" (Mackenzie 2014, 41).

Thinking about autonomy as a matter of degree evaluated relative to contextually appropriate thresholds helps alleviate the concern that the model of autonomy I have presented above is too demanding. It does this in two ways. First, by proposing that we need not maximize our performance of all the meso-level functions, it allows that autonomy is still compatible with imperfect micro-level skills and attitudes. This is consistent with our common-sense understanding of people as having a variety of strengths and weaknesses, and as rarely able to achieve and consistently maintain a maximal level of proficiency of any individual skill, let alone several skills at once as the multi-level conception of autonomy might be taken to require.

Second, seeing autonomy as a matter of degree invites the possibility that to be autonomous, people must achieve a particular point on the spectrum of degrees of autonomy attainment in

order to be attributed that status and benefits that are associated with being autonomous in society. But, autonomy is further removed from being an all-or-nothing property by accepting that in different contexts different functions, or levels of functioning, will be important, and so the threshold for being recognized as autonomous will vary in relation to the situational demands and implications. As a result of identifying these features of the model we are more likely to believe that people are generally autonomous because they are not expected to be perfect and because in mundane cases the threshold for qualifying as autonomous is relatively low.

However, there are remaining worries that suggest that conceiving of autonomy as a matter of degree is insufficient to show that autonomy is widely enjoyed. A first potential problem is that while understanding that the status of being autonomous socially requires only rising above an often not terribly high threshold suggests that many people should be capable of qualifying as autonomous, we might equally be concerned that autonomy is quite precarious in that many people might only function at the level just above the threshold and could drop below it if they experience certain changes. An account of how people can withstand or resist losses in autonomy resulting from other changes is therefore needed. A second source of concern is that those whose skills and attitudes meet lower thresholds of autonomy, may still not have sufficiently expert autonomy functioning to meet the higher thresholds that might be demanded by the more complex aspects of social life and political participation. I will attempt to address concerns of this nature in the sections that follow. The first way I will do so is by drawing the reader's attention to one further mechanism, *automaticity*, that is available to people to reduce the cognitive burden of exercising autonomy. In this way, this mechanism should help temper expectations about the mental resources required to meet the conditions for autonomy. It is also a means by which autonomy can be exercised by those with deficits in ways that help them avoid

falling below relevant thresholds, or meet thresholds that would otherwise be just beyond their reach.

ii.   Automaticity

As we saw in Chapter 2, one of the potential ways that autonomy could be so overly demanding as to be implausible relates to the mental demands of introspection. The concern expressed at that point was that people would have to spend quite a lot of time occupied in introspection and deliberation, since it is so important for our decisions and actions to be in alignment with our authentically held self-defining attitudes. There was a concern that this prescription implied that for any action to be autonomous it had to be the immediate product of a process of deliberation, meaning that spontaneous but authentic action would not be possible. Concerns of this nature reveal at least three different problems for autonomy. The first two relate to whether actions spurred by unconscious forces could count as part of the exercise of autonomy. The first is that autonomy is overly cognitively demanding because it must be the result of conscious processes, and the second is that autonomy is incompatible with spontaneity. Third, we might worry that autonomy requires that we carry out deliberations in their entirety whenever we need to choose and act.

To address the broader concern about autonomy and introspection, I drew on the work of Diana Meyers to defend the consistency between autonomy and spontaneity by arguing that actions founded in habit, natural inclination, and unconscious "agentic powers" could still be reflections of an agent's exercise of autonomy. We might think, however, that these earlier worries about the continual need for deliberation and deliberate action to count as autonomous are exacerbated by the ways in which the micro-level skills and attitudes may be more or less effective in some situations than others, and thus open to being derailed. This suggests that use of these faculties is

bound to occasionally go wrong. My arguments in this sub-section are partly intended to respond to concerns that this makes the exercise of autonomy impracticable. In addition, however, I will argue that beyond merely being compatible with autonomy, if certain behaviours or choices can be made automatic through the conscious development of habits and the adoption of pre-commitments, then these will also count as the result of the exercise of autonomy (at least while they remain endorsed).

First, however, I would like to clarify my position on the use of habit and other deliberative shortcuts in exercising autonomy. I propose to do this by considering a preliminary objection to the reliance on mechanisms like heuristics and biases in the exercise of autonomy. There exists a considerable scientific literature that supports the belief that people's judgments and actions, for better or worse, are guided by heuristics and biases.[71] A recent review of 214 studies revealed evidence for at least 19 different types of biases and heuristics involved in decision making (Blumenthal-Barby and Krieger 2015, 5-6). Even if, as I advised in Chapter 1, we are cautious about applying the findings of empirical research because of the systemic reproducibility problems in psychology and related disciplines, we still can be relatively confident that judgement is influenced by at least some such mechanisms (Schwab 2016, 26). Our confidence on this point is based both on the advanced state of research in this area and, more importantly for our purposes, on our common-sense understanding of human psychology.[72] For example, without committing to the existence of particular biases only a little introspection should reveal the biases we have that either favour or disfavour certain people, or in the way that we consume

---

[71] Given that many of these heuristics and biases operate unconsciously in our deliberations, we might see this objection as an extension or relative of the situationist concerns about autonomy articulated in Chapter 1.

[72] The support of both is valuable for two reasons. First, while the research is advanced enough, according to Schwab, to be confidence that heuristics and biases are at work generally, there is insufficient support for individual heuristics and biases. Second, the evaluative standard I have set myself is tied to common-sense.

information. However, even such a weak and general claim about the influence of biases and heuristics on judgement raises natural concerns about to the extent to which they pose a threat to the possibility of autonomy (Meyers 1989, 173, Blumenthal-Barby 2016, 5). Indeed, J.S. Blumenthal-Barby has illustrated how particular biases can (among other things) undermine or distort people's ability to understand "the nature of the action [they are choosing] and the foreseeable consequences" (Blumenthal-Barby 2016, 9), cause people to abandon their authentic intentions (Blumenthal-Barby 2016, 10), or lead people to feel alienated from their actions.

In reflecting on the conditions of autonomy that I have outlined above, it should be relatively clear that these influences are at least potentially problems from the perspective of autonomy. So, it would seem that there are good reasons for doubting that relying on heuristics and biases to lighten our cognitive load can be conducive for autonomy since they may lead to erroneous or fallacious conclusions or interfere with our motivations. There are three responses to be made to those who, like Blumenthal-Barby, are concerned about the prospects and implications of more easily exercising autonomy through reliance on biases, heuristics, and similar mechanisms. The first is that I'm not claiming that we should ignore the potentially autonomy-undermining aspects of these cognitive mechanisms. In some ways, taking on board empirical findings about cognitive biases should lead us to do much as John Doris and Eddy Nahmias suggested in relation to situationist insights, which was to incorporate these warnings into our self-reflections and deliberative practices to mitigate their impacts (Doris 1998, 517, Nahmias 2007, 12). Second, even Blumenthal-Barby notes that not all "cognitive biases and heuristics are detrimental to good… decision making" (Blumenthal-Barby and Krieger 2015, 550). She acknowledges that some promise to speed decision making while not significantly undermining accuracy. So, there is room to accept that the role of biases and heuristics in judgement is mixed.

And finally, it is worth noting that I will only be defending the role of particular kinds of heuristics and decision making shortcuts in the exercise of autonomy. Specifically, the kinds of automaticity that I will propose can be incorporated into the exercise of autonomy are ones that are intentionally cultivated. I'm not suggesting that just any heuristics can play this role, and would resist including the heuristics and biases and other automatic responses people simply end up possessing without subjecting them to careful scrutiny.

The difference between forms of automaticity that are part of autonomy and those that are not can be illustrated by pointing to the difference between two kinds of proposals advanced by Diana Meyers. The less plausible, in my view, is her proposal that the untrained functions of the self-as-embodied count as an exercise of autonomy. In "Decentralizing Autonomy: Five Faces of Selfhood", Meyers considers five conceptions of the self, claiming each "captures a significant dimension of selfhood—of what it's like and what it means to be a human subject" (Meyers 2005, 31). She thinks these include conceptions of the self that are unitary, social, relational, divided, and embodied (Meyers 2005, 29). For Meyers, the embodied self is a dimension of selfhood that is importantly connected to people's physical self-image and capabilities, and the domain in which autonomy is affected by changes in health and bodily condition (Meyers 2005, 31). As part of her project of establishing that autonomous action is not always reliant on rational reflection and deliberation— skills she conceives of as associated with the unitary, mentalistic self— she proposes that people can be deemed to be exercising autonomy even when their body has "taken over" (Meyers 2005, 33). Meyers relates a personal account of a hike during which she fell and broke her wrists. After having decided that the best course of action was to continue downwards (rather than back to a facility at the top of the mountain), she claims that without engaging in further deliberation her body acted so as to prevent further injury and released

adrenaline to facilitate her progress towards the bottom by making her feel energized, while reducing her pain and fear.[73] She urges her reader to "entertain the possibility … that [her] self-as-embodied… orchestrated an autonomous descent" in part because she did not resist her body's functioning at the time and retrospectively endorses it. That is, though she did not choose (or doesn't see herself as having chosen) to act the way she did, the lack of alienation she experiences relating to this episode and the recognition of her embodied self as part of her identity suggests that her actions are not heteronomous, but are what she would call "autonomy-preserving" (Meyers 2005, 35-6).

The problem with this position as I see it is that, as Meyers acknowledges, on her description of affairs she was simply lucky that her body responded in this way. Even if we grant that her body produced an unreflective but effective response to the situation, and that her embodied-self is part of her identity, it seems to me that the fact that her body's response was effective and not something that she regretted or feels alienated from could well be a matter of chance. She admits that little of her body's response could be attributed to "learned mountaineering skills", and so she claims that "the idea of a trained, adept body isn't quite the point" (Meyers 2005, 34). However, this says to me that her body's response could just as well have been to lose consciousness, go into shock, or any number of other ineffective, potentially embarrassing or life-threatening options.[74] Taking her account at face-value, it seems more like an instance of

---

[73] There is certainly room to question whether Meyers's interpretation of what happened in this situation is accurate, especially whether she *really* engaged in the descent totally unreflectingly. Even Meyers herself has said that autonomous people will be alert to problems during their automatic or habitual exercise of autonomy. So, there is reason to think (thanks again to R. Myers here) that she would have reasserted control in the presence of appropriate signals that her "exercise of [autonomy] competency" was deficient (Meyers 1989, 86). But, for the purposes of the following discussion I will not contest her account on these grounds. In my estimation it is more important to consider philosophical dimensions of what she is trying to convey through the example than to get overly exercised about the details of the account.
[74] I myself have had a similar experience, I broke my ankle walking alone in the Peak District and would rather say that I fought fatigue, dehydration, and shock all the way back down to the village, on to the train, and through

fortuitous co-incidence between bodily reaction and retrospective authenticity. She attempts to reconcile the extent to which her response was a matter of luck with her autonomy by claiming that she learned about herself— in the words of my multi-level theory performed the function of self-discovery— through her experience on the mountain (Meyers 2005, 43). She views this as an instance of self-discovery through action, which she does not think needs to be intentional (Meyers 2005, 48). As such, she believes that her experience on the mountain should be interpreted as a successful exercise of unconscious agentic skills in a way that meets both her requirements that autonomy depends on the use of agentic skills and should produce a sense of wholeness (Meyers 2005, 48-9).

Part of the reason some might object to Meyers's hiking example is that her presentation makes it seem like getting down the hill was an anomaly. Her explanation that her body "took over" suggests that there is no psychological explanation for how she made it down the mountainside. This suggests that Meyers is trying to characterize mere behaviour not only as action but as action stemming from the exercise of autonomy, which would be problematic indeed. But, Meyers's gesture to the importance of a sense of wholeness suggests that this reading of her argument is not adequate and hints at a more plausible way of understanding her account. At the end of the paper Meyers warns against reading her different conceptions of the self as compartmentalized. Rather, she sees them all as accommodated within a "a single self-conception" (Meyers 2005, 50). Thus, her focus on the self-as-embodied in articulating the hiking example conceals the possibility that her descent from the mountain, while not

---

Manchester on the way back to my lodgings. In my case, I perceived every step of the return trip as calculated move, which I am not sure makes me more or less autonomous that Meyers.

consciously relying on the skills, strengths, or attitudes that manifest consciously in other parts of her life, can be understood through them.

It seems necessary that Meyers would need to be courageous (and perhaps also resolute, persevering, prudent, and many other things) in ways that are transparent to her in other contexts for the requisite sense of wholeness to manifest itself. If not, then being unable to account for this inner strength should have left her feeling alienated in precisely the way she thinks is inconsistent with autonomy.[75] We might, therefore, be inclined to think that Meyers has at least somewhat consciously practiced, developed, or endorsed relevant self-defining attitudes in other areas of her life and her hiking experience is perhaps best interpreted as learning how those attitudes translate to a novel situation and perhaps shed light on their extent. It should be no surprise that self-defining attitudes, particularly in the form of values about the kind of person one wants to be (patient, compassionate, courageous, etc.) might lead to different and perhaps unanticipated responses in different situations. Moreover, there may be several different responses that are consistent with one's authentically endorsed self-defining attitudes in any given situation. Meyers's insistence on the importance of a sense of wholeness suggests that the circumstances she is describing would be best characterized as an instance of unconscious action in accordance with somewhat abstract self-defining attitudes, which happened to take a particular form on this occasion. It's not that she was trying to be courageous and persevering, it's that she *is* courageous and persevering and this was manifested in her automatic response to her situation.

---

[75] We might imagine an alternate persona called Beyers who understands herself to be unathletic, faint of heart, and easily discouraged who, upon making it down the mountain through the agency of her body cannot account for it at all, and who might feel alienated or unnerved by the result with which Meyers is readily able to reconcile with her self-conception.

To my mind, this brings the hiking example more into line with other, more plausible examples of autonomous action carried out by the self-as-embodied she outlines elsewhere in the paper. I would contend that these examples are more plausible because they all involve the idea of a body trained to respond in particular ways. For example, she lists the ability to swim as a potentially life-saving bodily response to a boating accident (Meyers 2005, 34), as well as the use of self-defence techniques for responding to bodily threats (Meyers 2005, 40). The objection that might be made to the way Meyers articulates her hiking example is that it conflates, or makes impossible to distinguish, action that is part of the exercise of autonomy and *mere* behaviour or, in Meyers's words, "doings" (Meyers 2005, 36). Under the kinds of mentalistic conceptions of autonomy that Meyers is trying to reject, confusion of this nature is less likely since action autonomously undertaken tends to already be intentional. That is, the concern on more standard accounts of autonomy is to distinguish between "merely intentional action and action that is, in the fullest sense, the agent's own" (Benson 2005, 101). Meyers's proposal of the self-as-embodied changes the question to consider whether we can have the necessary kind of ownership over our (at least superficially) unintentional actions for them to also qualify as autonomous.

I would argue that the difference between mere doings and autonomous doings will not be articulated in the same way as the difference between mere doings and action. For example, we might think that what makes action distinct from behaviour is that action is connected somehow with our intentions or goals, whereas behaviour is unintentional. But, in the distinction of interest for us, both kinds of doings are not the immediate result of conscious deliberation or choice. So, in this sense they are both unconscious. As I've suggest above, one way in which actions can be attributed to the exercise of autonomy while not being the immediate product of conscious deliberation or choice is that they are a trained reaction based on endorsed self-defining attitudes.

As noted in footnote 73, there's no reason to think that such trained responses are unalterable or impervious to reflection, only that our way of monitoring these actions is more passive, involving alertness to signals (such as negative emotional responses) that suggest that a mismatch has emerged between our habits and our authentically held self-defining attitudes (Meyers 1989, 86).

Thus, the first kind of automatization that I have in mind as potentially part of the exercise of autonomy is habit or habituation. This is suggested by Meyers's examples of swimming and self-defence. The kinds of habit that toward the exercise of autonomy are those that are developed intentionally through practice and repetition. This is why instinctive responses to swim or break a hold can count as an exercise of autonomy even if the reaction in the heat of moment is too quick to be fully thought through. The values or goals that prompt the cultivation of these physical abilities are adopted long before the emergency in which they are deployed unconsciously. And, the nature of the reaction is not decided in the moment, but is rehearsed and ingrained through a deliberate training process. So, the habits I have in mind have an important relationship with the mental life of those who possess them, it's just that actions based on these habits are connected to deliberation and choice by processes of inculcation that occurred at some point in the past rather than being determined in the moment. My reading of Meyers's hiking example reveals two further notable features of such habits.

First, the habitual responses that are part of the exercise of autonomy need not only be bodily responses. Taking Meyers at her word that her mountaineering skills were not highly developed, I looked at the psychological traits that were suggested by her ability to continue her descent despite her injury. I proposed that, given that she was not startled by or alienated from her actions, she should be expected to be courageous, resolute, and persevering in other aspects of

her life, which is what makes her response both natural and intelligible to her. Consequently, the habits in question need not be physical in nature, they can equally be habits of mind.

Second, my reading of Meyers's account suggests that at least some habits can guide action in more than one context. We might imagine that Meyers's courage was cultivated in her professional career, where she had to learn to master fear and anxiety when speaking publicly or facing critics.[76] One interpretation of her experience on the mountain is that at least some of the mental habits and coping mechanisms that she employs regularly in her career subconsciously contributed to her ability to keep going. On the whole, however, I would not expect that it can be immediately assumed that people's habits are an extension of their autonomy. Those that are intentionally developed in response to authentically adopted self-defining attitudes are the most natural candidates for counting as part of autonomy. Conversely, we would expect that behavioural ticks, such as nail biting or pen clicking, would not satisfy this requirement. Consistent with the accounts of authenticity articulated earlier, it does seem that habits that one did not cultivate intentionally, but that one ended up with somehow *can* become authentic through endorsement. Thus, there are two pathways through which habits can become an extension of the exercise of autonomy.

I turn now to what I see as the second way in which autonomy can be exercised while by-passing some of the cognitive load associated with deliberation and choice. This is through the use of pre-commitments. The idea I have in mind here is importantly different from the use of commitment mechanisms, which usually involve reinforcing desired behaviour by connecting it with something desired or something feared (Milkman, Minson, & Volpp, 2013, pp. 283-284). It

---

[76] Or even defending her dissertation!

is also different from choice architecture in that this is usually a strategy used to manage the behaviour of others and my interest is in an individual's production of authentic action while reducing the associated cognitive burden. What I have in mind is that by adopting certain values or personal policies, one can restrict the number of options that need to be considered or deliberated among for the purpose of exercising autonomy. My claim is that in order to act autonomously, a person does not need to consider all the possible options and match them against their self-defining attitudes in every situation. Instead, personal policies can pre-emptively rule-out whole subsets of options based on prior commitments. Thus, given a set of authentic values, only certain categories of actions are relevant, while other categories can be dismissed out of hand.

We might expect that in many situations the filtering role of pre-commitments leaves people with only limited viable options to consider, and possibly only a single default response. This would also, in a somewhat more conscious way than habit, reduce the cognitive burden associated with exercising autonomy. For example, a vegetarian who goes to a restaurant does not necessarily consider all the menu items. There may be entire sections of the menu that can be written off immediately as irrelevant, not even to be considered: for instance, the section of a menu having to do with what is on the grill. Indeed, a vegetarian may find himself with only a single option given this pre-commitment—this would be unlucky from the perspective of culinary variety, but at the same time his cognitive burden would be minimized. This example illustrates how authentically held commitments made long in advance of any particular decision can constrain the options to be considered so much as to make the cognitive burden of any particular decision much lighter than an emphasis on the importance of deliberation would suggest. This is precisely

because pre-commitments make it unnecessary to do all of the weighing and evaluating that could be called for when choosing.

The automatization mechanisms I have considered in this chapter go some way toward illustrating why autonomy is not overly demanding. Conceiving of autonomy both as a matter of degree and at least partly realizable through automatic mechanisms promises to decrease the perceived demandingness of autonomy, however both these features continue to rely on a personal ability to exercise self-control. In the case of habits, we need to exercise enough control to practice the behaviour to the point where it becomes internalized, second nature. In the case of pre-commitments, self-control appears to be necessary in order to prevent oneself from being distracted by options and possibilities that are in conflict with one's pre-commitments. So, while acknowledging the availability of these mechanisms decreases the apparent demandingness of autonomy, there is still a reliance on personal psychological powers which some may lack. As such, it would appear that autonomy could still be quite precarious and for many potentially unattainable. In the next chapter, I will turn to a significantly different set of strategies and methods that are available to facilitate or enhance the exercise of autonomy. These should help complete the picture of why it is plausible that many people exercise sufficient autonomy to warrant social consideration, despite the deficits in their micro-level skills and attitudes

Chapter 5: Filling in the Picture

This chapter continues the project initiated in Chapter 4, which involves articulating the additional features of the micro-level of analysis that help defend against the charge that autonomy is impractical both because it relies on an *extensive* number of skills and attitudes, and because of the *complexity* of the relationships between the meso- and micro- levels of analysis describing the exercise of autonomy. The previous chapter considered how the idea of autonomy as a matter of degree and the potential role of automaticity in autonomy contribute to the plausibility of the idea that autonomy is widely practicable. These features were treated separately because they are already well recognized and because they rely on relatively intact personal psychological capacities, especially those related to self-control. This chapter takes up what I see as less well recognized ways of defending the practicability of autonomy that compensate for significant deficits in capacities at the micro-level of analysis. What should be becoming apparent to the reader— and should become increasingly so through the first section of this chapter— is that although the additional features I am explicating make the exercise of autonomy more complex, it is precisely this further complexity that contributes to upholding the plausibility of autonomy. The salutary role of *complexity* should become increasingly clear as I articulate how processes of reinforcement and substitution at the micro-level provide multiple routes to achieving the functioning described at the meso-level of analysis.

Substitution and Reinforcement

So far, the further aspects of the multi-level model I have identified as supporting the idea that autonomy is widely enjoyed have focused on the internal lives of would-be autonomous agents. My contributions in the previous chapter have made it clear that although we should expect people to be flawed, and to have less than perfect skills and attitudes at the micro-level, these deficits are not to be viewed as incompatible with autonomy. In the subsection defending the

idea that autonomy is a matter of degree, the flaws of human nature are reconciled with autonomy by establishing that people only need to reach a threshold level rather than mastery of the micro-level means in order to be counted as autonomous. In the subsection defending the role of authentic automatization of actions and responses in autonomy, I argued that psychological mechanisms could be used to reduce the cognitive burden of autonomy. And, I proposed that reliance on default responses grounded in authentically endorsed habits or pre-commitments could compensate for micro-level deficits. In this section, I will discuss the extra-personal mechanisms for contending with deficits in the micro-level skills and attitudes that are the basic precursors of autonomy. I will argue that we can draw on resources in our environment either to bolster our micro-level abilities, or as a substitute for a micro-level ability that is lacking or absent. Both can be ways of coming to function adequately at the meso-level despite micro-level deficits.

My defense of these alternative routes to exercising autonomy will resist the temptation to classify the elements that make up autonomy as strictly internal or external to the agent. Catriona Mackenzie appears to succumb to this temptation at least somewhat in her articulation of the three axes she takes to be involved in autonomy.[77] The axes she names are "self-determination, self-governance, and self-authorization" (Mackenzie 2014, 18). The items she includes in the self-governance and self-authorization axes are largely the same as the elements I take to be of interest at the micro-level of analysis. Self-governance picks out the "skills and capacities" involved in reflection, comprehension, introspection, decision-making, communication, etc. that are deployed when establishing one's authentic self-defining attitudes and acting according to

---

[77] Her other work suggests, however, that she does not in general endorse a strict distinction between internal and external functions.

them. Self-authorization describes the conditions for taking and exercising "normative authority" over one's values, commitments, reasons, and life more generally (Mackenzie 2014, 19). These conditions are closely associated with the possession of the self-regarding attitudes such as self-trust, self-respect, and self-worth, discussed both in the previous chapter and in greater detail in Chapter 2.

Mackenzie acknowledges that self-authorization is strongly social and relational. Our self-regarding attitudes are developed and reinforced by how we are treated by others. For instance, levels of self-authorization will vary depending on whether one's authority over one's own life is recognized socially (Mackenzie 2014, 19). However, she specifically classifies the various aspects of self-government as the "internal conditions for autonomy". It is my contention that the distinction suggested by Mackenzie's claims that (1) social forces are essential to self-authorization and (2) the skills involved in self-government are internal, is potentially misleading. At least in this characterization of self-governance, Mackenzie is not giving due consideration to the ways in which the skills and capacities at the micro-level of analysis can be augmented by reliance on extra-personal resources. That is, insofar as we would be inclined to propose that self-governance is a meso-level function[78], Mackenzie does not adequately acknowledge the full variety of means through which the *function* of self-governance may be achieved. Some of these means may include the internal capacities of an individual, but where a capacity or skill is diminished or lacking, functioning may equally be preserved through an agent's reliance on extra-personal means.

---

[78] As opposed to a synonym for autonomy as a whole.

While the conceptual model I have been describing has been to this point quite internalist, if not mentalistic, the important shift in this subsection is to capture the ways in which the skills and attitudes underlying autonomy are open to common, but sometimes insufficiently recognized, reinforcement or replacement by extra-personal resources. In particular, I will illustrate the opportunities that this creates for the exercise of autonomy in spite of our imperfections. The resources that provide the basis for enhancing the exercise of autonomy beyond what a person's individual abilities might suggest possible are available in what I will treat as two separate spheres: the relational and the environmental. By the relational sphere I mean to designate both the interpersonal relationships and the social institutions in which a person participates. By the environmental sphere I mean the physical environments in which people live. It will be noted that many of my examples relating to the physical environment involve the built environment, which is in turn a product of the social environment. As such, it may be argued that the environment should not be treated separately.

I would defend the separate treatment on two grounds. First, the ability to incorporate the environment into one's exercise of autonomy is underrecognized in contemporary theorizing about autonomy. Second, the kinds of strategies by which the environment can be used to enhance the functioning required to achieve higher levels of autonomy are substantially different from those than would be used socially. By the end of this subsection I hope to have presented a convincing case that, in combination with the acceptance of autonomy as a matter of degree and with the idea that autonomy can be exercised automatically, the facilitation of the exercise of autonomy by social and environmental means reinforces my claim that autonomy is not implausibly difficult. The functioning required to exercise autonomy can be achieved by different people in many different ways and using a variety of strategies.

An important project in contemporary thinking about autonomy is reconciling our ideas about autonomy with the fact of human sociality. In Chapter 2, I argued that interest in this question extends back at least to the seminal papers by Frankfurt and Dworkin. However, as discussed at that juncture, the relationship between autonomy and sociality has been an important point of focus in feminist thought. One prominent vein of feminist inquiry concerning autonomy has sought to "analyz[e] the way in which socialization and social relationships impede or enhance an agent's capacities for autonomy" (Mackenzie and Stoljar 2000, 22). In Chapter 2, I noted that this had been done in ways that highlighted causation and constitution. The more straightforward and intuitively compelling of the two, to my mind, is that autonomy is causally relational. To remind the reader, causal relationality expresses the view that autonomy is relational because the basic skills and attitudes that are deployed in the exercise of autonomy are developed for better or worse under specific historical and social circumstances. These circumstances influence the extent to which people possess them (Mackenzie and Stoljar 2000, 22). Like Nussbaum's basic capabilities, the skills and attitudes of interest at what I would call the micro-level of analysis, though basic, may not be hardwired, but instead require nurturing (Nussbaum 2011, 23).

It seems obviously true that the conditions under which one is raised will influence the level of autonomy one enjoys. Though I would contend that the autonomy-shaping effects of social conditions are relevant not only in childhood but continue to shape the skills and attitudes of the micro-level throughout adulthood. I take this to be true not only of, per Mackenzie's view, more outwardly oriented aspects of autonomy such as self-worth, which are highly "sensitive to others' attitudes toward the agent" (Benson 1994, 659), but also of aspects of autonomy that would fall under into what Mackenzie has framed as the "internal" conditions of autonomy such

as imagination (Mackenzie 2000, 125) or the ability to perceive the world for what it is (Wolf 1989, 145). Just as we would expect that being raised by people who don't recognize your independent value as a person undermines your self-worth, so being raised under oppressive cultural images or with a skewed evaluative perspective can distort or fetter one's abilities to imagine possibilities or interpret the world. Consistent with my thought that social conditions shape autonomy throughout one's life, it is worth acknowledging that even if one was not raised in oppressive circumstances, if one's social circumstances were to become oppressive one's personal micro-level means could be rendered less effective in similar ways.[79]

The second broad relational position on autonomy is constitutive autonomy. This view goes further than the causal view in that it claims that autonomy is possible only through relationships, either because the capacity of autonomy or because agents themselves are socially constituted (Mackenzie and Stoljar 2000, 22). On this view, the level of autonomy a person is capable of is not only the product of social forces that affect their development throughout their lives, but also composed of social dynamics in which they are enmeshed. While I fully accept that autonomy is causally relational, I think the kind of reinforcement and substitution for autonomy skills I am defending also shares some features of constitutive accounts of relational autonomy. I will argue that people may rely to a greater or lesser degree on other people with whom they have personal relationships, or on institutions, to compensate for deficits they face in the exercise of autonomy in a way that possesses features of both causal and constitutive relationality.

 Holger Baumann has offered a useful analysis of the differences between the role that social factors play in causal and constitutive positions on the relationality of autonomy.

---

[79] Thanks to R. Myers for this point.

According to him, on a causal account of relationality social factors are involved in "the development and on-going exercise of autonomy" (Baumann 2008, 447). They serve as background conditions or contribute to autonomy (Baumann 2008, 447), and are therefore are relevant to autonomy insofar as they have an impact on people's capacities involved in autonomy (Baumann 2008, 451). Importantly, he contends that on a causal account, autonomy is still understood in an individualized fashion relating to the presence of particular psychological states or abilities (Baumann 2008, 447). This means that "the question whether a person is autonomous can be fully answered with reference to her internal psychological states" (Baumann 2008, 447).

Examples of causally relational autonomy are quite common. For instance, self-direction— understood as the ability to act in accordance with one's authentic values (Meyers 1989, 43)— and self-authorization are arguably the autonomy functions that are most vulnerable to being defeated by social factors, and most open to social support. Self-direction and self-authorization are so significantly under the sway of social pressures because of how closely connected they are to social circumstances and practices. Our abilities to carry out plans based on our self-defining attitudes are highly dependent on having the social permission to do so. Social restrictions on how we move around in the world and on which institutions or cultural practices we are eligible to participate in play an important role in determining which goals or values we can pursue. The kinds of support we might obtain in our ventures will also be conditioned by these factors. So, while there may be a putatively internal motivational aspect to self-direction, it is also highly dependent on "having the freedom and opportunities to make and enact choices of practical import in one's life" (Mackenzie 2014, 18). Similarly, our attitudes about our own worth and competence, insofar as they are sensitive to the appraisals and recognition of others and of

institutions, are also readily influenced by social factors for the better or the worse[80] (Killmister 2013, 116). In these cases, social factors are merely influencing the ways and extent to which we are able to live according to our authentic values by creating constraints and freedoms or putting pressure on our self-defining and self-regarding attitudes.

The main feature of a constitutive account of relationality, according to Baumann, is that autonomy is understood to be partly constituted, in the sense of defined, by social factors or requirements (Baumann 2008, 447). On such a view, autonomy cannot be described without reference to the social environment. As such, social factors are not just background conditions, but they are related to autonomy in a much more intimate way.

In contrast to the examples of the ways that social factors can influence autonomy above, a constitutive account of autonomy might hold that someone can't be autonomous without living in social circumstances in which they are free to participate in and pursue the ventures of their choice or in which they have a respected social status. In reframing the issue in this way, it is not just the case that our autonomy is affected by social conditions, but that autonomy is dependent on presently enjoying particular social conditions.

My own view agrees with the causal stance in that it sees relational factors as playing a role in the development and ongoing exercise of autonomy. However, my view departs from causal accounts of relationality insofar as they contend (i) social factors only *impact* one's capacities and (ii) we can assess whether someone is autonomous based solely on their internal psychological states. It is my position that, similar to constitutive views of relationality, the

---

[80] Certainly, our autonomy can be undermined by having our worth insufficiently recognized. But, it's interesting to think that more in this case isn't always better in the sense that if one's ability to see the world for what it is and other capacities are undermined by boosterism, the result could also be bad for autonomy. Thanks to R. Myers for this point.

relationship between autonomy and social factors is far more intimate than a causal account seems to acknowledge. That is, through reinforcement and substitution social factors do not just affect autonomy, they become part of autonomy.

It is because of this that I do not think it is possible to answer questions about whether someone is autonomous only by referring to their internal states. If relationships or environmental factors are part of someone's autonomy, then they will need to be referred to in order to make a full assessment of the degree of autonomy someone is capable of. It is true however, that in understanding autonomy as the ability to live according to one's authentic values, I do not define autonomy in a way that shows these relationships to be conceptually necessary. At the end of the chapter I will explore the sense in which my view might be understood as a form of constitutive autonomy. But, first I will articulate what I take reinforcement and substitution to involve as ways in which autonomy is exercised *through* social relationships. In conformity with the common-sense standard I have adopted, the examples I will provide involve relatively commonplace practices both in individual relationships and in relationships between people and institutions.

Recall the case of Dolores from the previous chapter. Dolores's situation, experiencing reduced bodily control resulting in loss of self-trust, is not dissimilar from the experiences of older adults with whom I have spoken in the service of other research. Dolores's decreased mobility and her dependence on an oxygen tank suggest that she would have difficulty accomplishing even quotidian tasks related to her values. It might be important, for example, to Dolores to keep a clean house because she associates this with a kind of dignity and sees it as a reflection on her taste and the pride she takes in herself. However, with the challenges she faces, she might find it difficult or impossible to clean her house (or to clean it to her satisfaction). But, Dolores does not

have to give up on the values that she thinks cleanliness represents if she can rely on others to help realize these values. Of course, if she has children they might be relied on to help her be self-directing by helping her clean. But, we might worry that this means only older adults with children can have their exercise of autonomy preserved or enhanced in this way. This worry can be met by considering one of the more notable features of my interviews with older adults who access home and community support services.

During my discussions with older adults, I found that it was not uncommon for them to report banding together with other people their age, either friends or neighbours, for mutual support. In Dolores's case she might have a friend who is more able-bodied who can help her clean, but whom she can also assist in some other way (with her company, pet sitting, etc.). Indeed, sometimes it seemed like this model was preferred since many older adults expressed reluctance towards, or regret in, depending on their children. Mutual dependence among people living in similar circumstances seemed not to be tinged with these same feelings. This could be because mutually supportive people are on more equal footing, or perhaps because peer-to-peer relationships do not represent the same reversal of child-parent roles. Reliance on children and friends to aid in one's self-realization is just one example of how people can leverage interpersonal relationships in their efforts to live according to their values.

It is possible to further develop Dolores's case to also illustrate how people might depend on institutions to be self-realizing. In the original description of her case, it was noted that Dolores did not trust herself to bathe on her own because she feared falling. If Dolores feels strongly about having a clean house, as depicted above, it is reasonable to think that she also strongly values being clean herself. So, not feeling able to bathe would be distressing and would represent a not insignificant blow to her autonomy—especially her ability to live according to her values.

One way in which Dolores might rely on institutions for help in remedying this is in obtaining access to a personal support worker (PSW) through a social services program for older adults. One of the personal support worker's responsibilities could conceivably be to assist Dolores in bathing, thus allowing Dolores to be self-directing despite her lack of trust in her own abilities to complete the tasks that this would entail on her own.

There are some drawbacks associated with relying on institutional support, not the least of which are (1) PSWs are not available everyday, and (2) PSWs are scheduled by their employers, and not necessarily at times that are preferred by the clients. As such, Dolores's bathing schedule would likely be dictated by the availability of her PSW, meaning she may not be able to bathe with the frequency or at the time she wants. For some older adults a superior alternative would be to live in a retirement residence setting, where there are staff on hand at all hours to respond to residents' needs. In this setting Dolores could decide when she wanted to bathe and could test the full range of her physical abilities, knowing that she had a safety net she could call on if she pushed herself too far, or felt unsafe. While having a safety net involves in part compensating for her lack of self-trust by placing trust in someone else, having someone else to rely on also promises to increase Dolores's self-trust, or at least to ensure that she lacks trust only within the range that is justified by her physical limitations. That is, by reducing the severity of the consequences of misplaced self-trust Dolores may find that she is able to do more than she thought because she may be less likely to overcompensate by becoming overly cautious because of her disability. Having this kind of safety net conceivably helps allay the fear that could contribute to a greater than necessary degree of self-restriction.

Although this illustration focuses on Dolores and her particular experiences, I hope that the reader can imagine how analogous analyses would be available for many other circumstances in

which people experience deficits in meso-level functions such as self-direction and self-authorization. Other people can participate in our efforts to be self-directing by doing all or part of what we want to have done. Institutions can make people available to fill this assistive role, or institutions can create opportunities for autonomy to be exercised in a setting in which deficits are compensated for, by, for example, ensuring a person can be safe in exercising autonomy to the fullest extent she is able despite her limitations. As I said earlier, however, we needn't restrict the prospects of relational support for autonomy to the aspects of autonomy that might be viewed, through Mackenzie's axes of autonomy for instance, as the social aspects of autonomy. The next few paragraphs will illustrate how relational resources can compensate for aspects of autonomy that some might think are more internal, which should also convey why it is not really appropriate to attempt to divide the skills and attitudes of the micro-level into internal and external categories. The examples should show how these too are open to relational reinforcement or substitution.

Any number of examples can be given of how our engagement with other people can enhance our performance of the skills associated with the micro-level of analysis. Some useful examples of how relationships with confidants can help sustain the motivation needed to realize self-direction have been put forward by Meyers. She recounts how she had to change her diet in order to accommodate a new health condition. She credits her friends and spouse with contributing to her ability to motivate herself to eat within the guidelines of her new diet by also refraining from themselves ordering foods she could not eat and listening to her complaints about her new circumstances (Meyers 2005, 32). She claims they helped her overcome her resistance to change and helped her avoid temptation. But, Meyers goes further claiming that even acquaintances helped her maintain her regime. By informing them of her situation, she developed a sort of

context of social expectation that "pre-empt[ed her] occasional renegade impulses" (Meyers 2005, 33). Because of the way she, admittedly somewhat unwittingly, extended her intentions into her social relationships she found it easier to resist her desire to consume things that would be harmful.

Other commonplace examples can be constructed for other micro-level skills. Take for instance how imagination and deliberation can be enhanced relationally. When faced with a difficult choice or a tricky social situation, we may come to feel that our own efforts to arrive at an optimal solution or course of action yield unsatisfactory results.[81] Even if our own reflections favour a single option, we may be left feeling that it is still unsatisfactory. Consulting a friend can be helpful in this kind of scenario. In explaining the problem to them, we may have to think more precisely about what is at issue, in explaining our own thoughts about the solution to them we have to take a more objective stance on our own beliefs and conclusions, and our friend may help us not only by pointing out flaws in our reasoning or approach, but by helping us imagine options that we may have had difficulty conceiving of on our own. In coming to a better solution through the help of a friend and in acting on it, our autonomy should not be assumed to be undermined by the spectre of heteronomy. We still endorsed and adopted the resolution for ourselves, after all. Rather, our autonomy has been enhanced because by leaning on our friend's deliberative and imaginative capacities, we enhanced our own performance in these areas, and discovered a course of action that we could more readily and authentically adopt.

In thinking about how people can enhance their autonomy through institutions, I suggest we return to Egon's case, described above. We saw that Egon was not able to carry out, or act

---

[81] These comments can apply equally to deliberating about what values we should adopt quite separately from how to act on the values we do endorse. Thanks R. Myers for this point.

according to, his authentically endorsed values and goals. But, in contrast to Dolores, this is not because of a physical inability, it is instead—so far as current scientific research can establish—the result of a breakdown or deficit in neuropsychological function, though many details are still contested (see for example Curatolo, D'Agait, and Moavero 2010, and Biederman et al. 2004). Some theorize that people with ADHD experience deficits in executive functioning that interfere with working memory, various aspects of introspection and self-reflection, internal motivation and problem solving (Barkley, Murphy and Fischer 2008, 460). Whatever the specifics, it appears that Egon's difficulties with self-direction have a significantly different basis than those experienced by Dolores. Indeed, the areas in which Egon experiences difficulties are not obviously open to social influence in the ways that self-trust and self-direction are.

To manage his symptoms, Egon will likely need to employ special techniques and require accommodations at work (Barkley, Murphy and Fischer 2008, 455). There may be some strategies that Egon can use privately to perform better. But, he may also require changes to his social environment, such as shortening the time between performance and feedback/consequences, setting up external and regular sources of motivation, and breaking up large projects into smaller tasks each with their short-term deadline (Barkley, Murphy and Fischer 2008, 461, 463). Thus, managing his condition may, for Egon, rely on the complicity of his supervisors and a degree of institutional flexibility potentially indefinitely in order to accommodate the measures that will help him be more effective at acting on his self-defining attitudes. Barkley, Murphy and Fischer see this process as tantamount to "designing a prosthetic social environment that enables the individual to cope better with and compensate for the disorder" (Barkley, Murphy and Fischer 2008, 465). This is notably not a situation in which abilities are being reinforced by the assistance or surrogacy of others, rather this is a situation in

which the social environment is being reconfigured to compensate for Egon's deficits in executive functioning. And yet, if Egon can perform better at work in consequence, I would be inclined to say that his ability to exercise autonomy has been enhanced.

ii.   Environmental Forms of Substitution and Reinforcement

Finally, I will turn to what I take to be the least well recognized way in which limitations in the skills and attitudes that are operative at the micro-level of analysis can be reinforced or compensated for through reliance on external means. The means I have in mind, as stated above, are not directly social. Instead they are associated with the physical environment. As I see it, there are at least two ways in which the environment can be relied upon to facilitate the exercise of autonomy. The first is modification and the second is navigation.

There are numerous ways to modify our environment so as to support the deployment of the micro-level skills that are the basis for autonomy. For instance, we might bring things we like into contexts we dislike. Those who dislike exercising might benefit from leaving a favourite book in their locker at the gym so that they only have the opportunity to read it while exercising. The pairing the book with exercise presumably helps increase motivation to exercise.[82] Or, someone like Dolores might benefit from putting up grab bars and constructing gentle ramps over any steps in her apartment. Adding these features to her residence does not ameliorate the bodily bases of the mobility challenges that are causing her to lose trust in herself. But, it does provide her with a new way to save herself from (or prevent) falls, which might increase her sense that she is still competent to live alone.

---

[82] This is based on Katherine Milkman's work on temptation bundling.

Importantly, someone like Egon may need to implement not only changes to the institutional structures he is part of, but to modify his working environment by posting signs with procedures and instructions in his visual field, playing recording of procedures while working, visually representing time and sources of motivation in his workspace (Barkley, Murphy and Fischer 2008, 463). All of these measures are supposed to overcome sources of distraction, which are hard to resist in the face of deficits in executive function, and to "cue [him] to do what he… knows [he should do]" (Barkley, Murphy and Fischer 2008, 463). While Egon's case might be extreme, we should be able to recognize in it aspects of commonplace strategies for avoiding distraction and increasing focus. The creation of lists of tasks and sub-tasks and the use of timers or time limits can be effectively employed by nearly anyone who is finding it difficult to stay motivated and focused. These examples reveal some of the roles environmental modification can play in the exercise of autonomy—it can reinforce our existing skills, it can provide us new ways of acting that may have implications for the exercise of autonomy, and it can compensate for skills that we are lacking.

Using environmental navigation to facilitate autonomy shares the situationist intuition that sometimes the most effective way to control our behaviour is to avoid situations in which our motivational effectiveness or self-control is likely to falter (Doris 1998, 517, Fujita 2011, 355). My proposal is that when our micro-level skills associated with self-control are lacking, ensuring that we act in accordance with our values by taking care in how we navigate through space is a perfectly valid measure in the exercise of autonomy. We might choose to avoid specific places, like when an alcoholic avoids a bar by crossing the street (Foddy and Savulescu 2010a, 13). Or, we might choose to avoid certain situations in which our ability to resist our less admirable inclinations or temptations might be overcome. For instance, we might avoid competitive

situations if we know we tend to be a poor loser (or winner). While employing environmental navigation as a means of exercising autonomy might seem a bit far-fetched, I would argue that it is actually more commonplace than we might recognize.

For example, until recently the advice for diet and nutrition conscious grocery shoppers was to *shop the periphery*.[83] That is, at least in North America, a common way of organizing grocery stores has been to put the fresh and whole foods (fresh baked goods, meats, dairy, and produce) around the outside of the store, with rows of processed foods occupying the interior. Those who shopped the periphery were using the process of navigating the physical space of the grocery store in order to reinforce their desire to eat a healthy unprocessed diet, cutting off avenues by which they would be undermined by the availability of alluring processed goodies. So, while we are not all using navigation of the environment as a way to compensate for lack of self-control resulting from addiction as in the case of an alcoholic, it is well within reason to believe that this strategy can be used as a way of also resisting more mundane temptations.

An important difference between what I am proposing and the situationist stance I rejected in Chapter 1, is that I am viewing the navigation of physical space in the aim of reinforcing the exercise of autonomy as a compensatory measure we can use when—as a matter of exception--our skills involved in sustaining motivation or exerting self-control are deficient. By contrast, a situationist in the tradition of John Doris will claim that we are in general wrong to think that our cognitive abilities can be relied on to ensure that our behaviour reflects our values in the moment (Doris 1998, 516). This might be a way of saying that our capacities for self-control and motivation are so highly vulnerable to situational pressures (Doris 1998, 510) that by default we

---

[83] I have the sense that grocery chains have now caught on to this and have adapted.

should avoid environments in which there is a temptation to act poorly (Doris 1998, 517). A further difference between my view and the situationist position is that, while I have to this point only offered examples associated with motivational effectiveness and self-control, this should not be taken to mean that these are the only micro-level skills that can benefit from relying environmental navigation for reinforcement or substitution. For instance, sometimes we use the process of retracing our steps to improve our ability to remember where we've put things. This kind of example could be extended and brought to bear on the exercise of autonomy.

The examples of relational and environmental reinforcement and substitution of micro-level skills and attitudes I have employed above are intended to represent real-life practices that are easily recognizable, if not common. Because they have been drawn from practices and situations that are part of daily life, I contend that they are consistent with the common-sense psychological outlook I have been employing. That is, even though the reader may not have spontaneously listed substitution and reinforcement as aspects of the exercise of autonomy, I hope their recognizability as part of normal behaviour appeals to perhaps more implicit elements of the reader's common-sense psychological framework. The same could be said for automaticity, to the extent that this not a regular part of our common-sense account of the exercise of autonomy. With a clear understanding of the potential alternative strategies for exercising autonomy, we are now ready to re-engage with the questions of complexity and practicability in the context of autonomy.

The Full Picture

The previous chapter, as well as the first half of the present chapter, have been dedicated to showing how, in addition to skills and attitudes, agents have at their disposal several further means by which to achieve the meso-level functions that are required for autonomy. These take

the form of habits, other people, institutions, and even the physical environment. The availability of these additional resources by which greater levels of meso-level functioning can be attained presents the possibility that autonomy in any given person may be, at its most basic level, grounded in a unique set of skills, attitudes, relationships, and environmental features. To return to the partonomic language I employed in part of Chapter 3, autonomy ends up looking like quite a special kind of thing. Specifically, given the account of the exercise of autonomy I have given in the preceding chapters, different instances of autonomy should be recognizable as autonomy even though they have different parts. In the case of autonomy, identity is retained despite not simply numeric, but substantive, differences in autonomy's basic parts from one person, or one instance, to the next. That is, many people may be able to determine what is important to them and live accordingly. However, what this will involve will vary from person to person not only in terms of values and goals, but also in terms of the underlying means that make the authentic selection and pursuit of their self-defining attitudes possible.

With my theory of autonomy now fully explicated, in this final section of the chapter— and of the dissertation as a whole— I am now in a position to respond to lingering objections and to finish spelling out the implications of my view. I will begin this section by addressing some potential objections and concerns about (1) identifying the necessary and sufficient conditions for autonomy, (2) the existence of hard cases, and (3) the compatibility of habit, reinforcement and substitution with autonomy. I will conclude this work with a discussion of some contributions I see this project as making, but which have not yet been given adequate support. These have to do with how I take my view to (1) vindicate the practicability of autonomy, (2) contribute to our understanding of the senses in which autonomy might be constitutively

relational, and (3) provide resources for rethinking the way that we respond socially to those who are deemed to not be autonomous enough for particular purposes.

i. *Objections*

As promised, below I will respond to some potential concerns about the way that I have attempted to reconcile the complexity of autonomy with its practicability. First, I will speak briefly again about why I do not think explaining autonomy through the identification of necessary and sufficient conditions, perhaps characterized by the meso-level functions, is productive. Second, I will address the existence of hard cases. The third and final objection I will respond to claims that strategies such as automatization, reinforcement and substitution do not really contribute to autonomy. Piper (who we met in Chapter 1) thinks that it is undeniable that autonomy requires a kind of control (2016, 773) and he would likely doubt that the alternative strategies I have proposed qualify as control in the appropriate way.

To address the first question, relating to stating the necessary and sufficient conditions for autonomy, we need to recall what I said about the special nature of autonomy at the beginning of this section. At that juncture I noted that, on my account, autonomy appears to be the kind of thing which can retain identity despite different instances being ultimately composed of different constituents. For example, one person's autonomy may include their personal capacity for self-control, while others may substitute their relationships with others or their strategies for navigating the environment for this personal power in order to achieve much the same effect at the level of functioning. The fact that they exercise autonomy in different ways does not negate the fact that they are all exercising autonomy.

This has important implications for any remaining inclinations toward stating necessary and sufficient conditions for autonomy. As I noted earlier, I'm not convinced that an effort to

definitively articulate such conditions is altogether appropriate. But, some might argue that the meso-level functions are each necessary conditions for autonomy (and together sufficient for it), and question why I would continue to resist talking about autonomy in this way. But, the ability to reinforce or substitute capabilities and attitudes in the exercise of autonomy suggests why focusing on the meso-level functions in giving an account of autonomy remains inadequate. As I see it, such an account would be inadequate for two reasons. The first reason is that a meso-level analysis of autonomy is not sufficiently granular to get to the fundamental roots of autonomy deficits. Even if we know that someone's level of autonomy is low, for instance because they are not self-defining, it is only by proceeding with a micro-level analysis that we can determine the underlying cause of their reduced function in this area, whether it be cognitive impairment, adverse social conditions, or something else. The second reason that autonomy cannot be fully understood through meso-level functions is closely related to what I have attempted to demonstrate through the first half of this chapter, namely that the meso-level functions themselves may be achieved through many combinations of factors, of *means*, at the micro-level. As such, stating necessary and sufficient conditions for autonomy in the form of the meso-level functions would provide only a thin and incomplete understanding that would not account for how instances of the functions may differ from each other.

The second family of concerns some readers may have relates to how easy or difficult it might be to establish whether someone's reliance on extra-personal resources represents an overall enhancement or detriment to their autonomy. A preliminary version of this concern is that relying on extra-personal resources, as part of self-regulation for example, seems to remove the impetus for people to cultivate their personal capacities for resolve or self-control and is therefore less than desirable. In general terms, I do not think that this is a terribly pressing

concern. I don't know that there is any principled reason why we should think that internal means are better or more legitimate than external ones. So, as long as someone is capable of one kind or another their exercise of autonomy is equally effective and valid.

However, I acknowledge there could be cases where someone's external resources or mental-load decreasing strategies are not available, in which case we might see them as doomed to fail to function adequately, and therefore to fall short of the threshold for meeting the conditions required for autonomy. This is a legitimate concern. Relying on extra-personal resources to exercise autonomy involves enmeshing oneself in a web of connections between people, places, and things. Heavy reliance on this web may mean that without it one's autonomy is significantly diminished. For some people whose personal abilities are not, or cannot be, improved as a biproduct of their reliance on extra-personal resources this may just be the bare truth.

There are three responses to be made to those who find this concerning. The first is that although such people will only be capable of exercising sufficient autonomy under certain circumstances, this is true --to a perhaps less dramatic degree-- of all of us. The second is that it's difficult not to suspect that qualms of this nature are grounded in lingering traces of commitments to conceptions of autonomy on which it is equivalent to independence and self-sufficiency. But, such views have been shown to be flawed in ways that suggest that departure from them should not be cause for concern. The third response is that I suspect it is also possible that some people's personal capacities can be developed as a background phenomenon of relying on extra-personal resources. For example, an alcoholic who crosses the street to avoid drinking still has to exert the self-control to cross the street or otherwise follow their self-regulatory[84] strategies.

---

[84] To be defined below.

Consequently, people in this category may be able to maintain sufficient functioning to count as autonomous without the extra-personal resources on which they usually rely for at least some time in some contexts.

As I advocated in Chapter 1, our judgments about who is functioning sufficiently well to qualify as autonomous is highly context dependent. This opens the door to a related concern that there may be cases in which it is difficult to tell whether someone's reliance on other people specifically is promoting or undermining their autonomy. It could be that the person being relied on is covertly exercising undue influence, or that the dependence relationship is beneficial in some ways, but detrimental in others. Admittedly these are hard cases, but I don't think that the framework I am developing for thinking about autonomy is responsible for ruling on cases like these. We have seen that social relationships can have positive and negative influences on autonomy, and particular relationships may be a mix of both. This fact does not undermine my view overall because despite this possible ambivalence, in both ambivalent and straightforward cases there is reason to believe that people *can* draw others into their exercise of autonomy—to better or worse effect. However, the possibility of autonomy-related ambivalence in people's relationships does suggest that those with healthy and supportive personal relationships will be in a better position exercise autonomy relationally than those whose relationships do not possess these qualities.

Finally, the third problem that some might have with my approach to establishing that autonomy is not overly demanding is that they might be unwilling to accept that automaticity, reinforcement, and substitution can legitimately be considered parts of autonomy. This concern has been developed in two slightly different ways by Piper his recent paper "Achieving Autonomy". In his first articulation of the objection, Piper claims that exercising autonomy

"…typically involves bringing to bear a careful kind of control in a conscious and deliberate attempt to conform one's actions to the demands of one's authentic self-conception…" (Piper 2016, 791). With the exception of consciously-established habits, Piper thinks that actions are only rightly called autonomous when they involved "focused acts of will". He claims focused acts of will are deliberate, mindful, and involve the expenditure of effort in the moment to achieve the authentically held goal in question (Piper 2016, 774).

I have endorsed the possibility that each of my alternative mechanisms for exercising autonomy can play a role in self-control, either through maintaining motivation or resisting self-defeating actions. Yet, none of them would qualify as acts of will, and they all represent ways of reducing the amount of effort that an individual has to expend in the exercise of autonomy. Piper, however, is willing to accept that habits formed on the basis of authentic values represent an appropriate kind of control (Piper 2016, 775).[85] So the problem is not, as Piper's initial comments about occurrent effort and control would suggest, that he has an overly narrow conception of what it means to act for a reason or intentionally.

He is likely to take greater issue with the ideas of reinforcement and substitution. My contention would be therefore that Piper has an overly narrow conception of what constitutes self-control. Following Kentaro Fujita, I believe it is worth noting that what people tend to mean by 'self-control' is more than just the conscious effort to inhibit certain impulses (Fujita 2011, 352). Rather, self-control is often used to refer to a more general process for managing goals and ensuring they are met (Fujita 2011, 353). For clarity, we might want to call this broader idea of self-control 'self-regulation'. It seems like Piper is claiming that only actions that are the result

---

[85] Piper has other concerns about how reliable habits are for producing autonomous action given that people change over time and that different social conditions might demand different responses, but these are outside the scope of the present argument (Piper 2016, 790).

of the use of effortful self-control can count toward the exercise of autonomy. But, for the purposes of establishing what counts as autonomy our interest should be, as Piper correctly states, the ability "to act, live, or behave in accordance with one's authentic self" (Piper 2016, 773). The other processes that make up the larger category of self-regulation are undertaken to achieve the same result as effortful self-control in that they are directed toward meeting one's goals. As such, only counting effortful self-control as self-regulation is not justified. As parts of self-regulation, reinforcement and substitution's roles in the exercise of autonomy should not be discounted.

The second version of Piper's concern with at least some of the mechanisms I have proposed make autonomy less demanding is that their acceptability relies on conflating autonomy and authenticity (Piper 2016, 788). Piper thinks it is wrong to believe that autonomy amounts only to being one's self, or acting according to one's authentic values. He contends that people can act authentically "thoughtlessly, by mistake, or even as the result of coercion, deception, or manipulation" (Piper 2016, 789). He claims that the key difference between autonomy and authenticity is precisely the requirement of "deliberate control" in the first, but not the second. He proposes that we might want to conceive therefore of a difference between actions that are autonomous, and actions that are *autonomy-consistent*. The distinction is based on the presence or absence of the use of a careful, conscious control (Piper 2016, 790-1). Piper might be tempted to claim that actions produced through the use of reinforcement and substitution would count only as autonomy-consistent, rather than as fully-fledged instances of the exercise of autonomy because they don't involve conscious control.

My response to this charge would be much as above. An excessive focus on effortful self-control wrongly ignores the other mechanisms by which we can achieve our ends—it is blind to the

185

broader picture of self-regulation. In addition, it would seem to me that substitution and reinforcement, while they reduce the need for occurrent effortful self-control in acting on our authentic values and beliefs, are importantly different from the other phenomena that he says fall under the scope of being autonomy-consistent. Off-loading the cognitive burden of exercising self-control on to our relationships or on to the way our environment is structured is not at all the same as acting authentically by mistake or as the result of coercion. The difference, to my mind, is that the other kinds of autonomy-consistent actions do not represent ways in which people seek to act according to their values. Substitution and reinforcement still have an important relationship to the process of striving toward living in a way consistent with one's self-defining attitudes. This aspect is not shared by the other kinds of autonomy-consistent actions Piper describes. In his view all that makes them autonomous is the consistency between the outcome and their values, but not the process by which these outcomes were produced.

ii.   *Contributions*

While some of the contributions I had promised to make, such as providing a richer conception of autonomy thresholds and reconciling coherence with *extensiveness*, have been well developed in earlier chapters, there are others that still require explicit discussion in order to convey the full significance of my view to the reader. These relate to questions around practicability, constitutive relationality, and the socio-political implications of my view more generally. I will address each of these points in turn.

First, it is because of the sheer variety of different ways in which meso-level functioning can be achieved that I do not think that autonomy is implausibly impracticable, even if it is complex. At the beginning of Chapter 4, I articulated the worry that the fact that autonomy appears to require the cooperation of so many different capacities and attitudes, presents equally if not more ways

in which the exercise of autonomy can go wrong. I noted that a common-sense appreciation of human fallibility will support the idea that most of us can be expected to have at least some level of impairment related to one or more of the micro-level skills and attitudes by which autonomy can ultimately be explained.

The sections defending the idea that autonomy is a matter of degree and articulating the mechanisms available to people to overcome and circumvent the deficits that would place them at relatively low levels of functioning in the exercise of autonomy, are a response to this concern. They tell us that there are ways that satisfactory levels of functioning can be achieved even in the face of imperfect micro-level skills and attitudes. The idea of autonomy as a matter of degree does away with the need for perfection altogether. And, the acknowledgement that meso-level functions can be achieved at least partly through automatization and extra-personal means provides the basis for claiming that even those with significant deficits in their personal level of attainment relative to one or more skills and attitudes are not doomed by their individual limitations. This is because most skills and attitudes are open to at least some degree of external support. Through substitution and reinforcement it is possible to avoid failures in meso-level functions and therefore to better exercise autonomy.

It is only in fully appreciating the complexity of how the micro-level operates that we can see how the multi-level model of autonomy is not overly demanding. It is because the micro-level comprises the broadly defined *means* that underlie meso-level functions, including the means associated with automaticity, reinforcement, and substitution, that our worries about individuals' weaknesses can be assuaged.

Second, the features of reinforcement and substitution in the account of autonomy I have advanced help describe the way in which autonomy is relational. Earlier in this chapter, I

grappled with the commitments held by causal and constitutive forms of relational autonomy. I found that my view shared features with both camps, but also lacked features from both. Although some might see my view as just a variety of causal relationality in which social factors are extremely deeply embedded in and widely spread through the exercise of autonomy, I rejected this interpretation because it is not clear that the intimate relationship between autonomy and social factors on my account is consistent with the idea that autonomy can be assessed purely on the basis of someone's internal psychological states. At the same time, my view is not constitutive in a conceptual sense—social factors are not invoked in the idea of deciding or discovering for oneself what's important and living accordingly.

To help clarify the relationship between my view and constitutively relational autonomy, it might be useful to note the idea that autonomy is constitutively relational derives from feminist conceptions of the self as constitutively relational (Christman 2004, 144, Mackenzie and Stoljar 2000, 22). This claim about the constitutively relational nature of selves, however, is shorthand for a number of different claims about the ways in which the selves are determined and constituted by society (Barclay 2000, 61). Being constitutively relational may be taken to mean that our self-defining attitudes—values, goals, beliefs-- are determined by our social context, that our ends as individuals are all shared ends with our communities, or that there are certain relationships or aspects of our social contexts that define and are essentially part of us (Barclay 2000, 61, 67, Christman 2004, 144).

What it might mean for autonomy therefore to also be constitutively relational presumably relates to at least one of these various ways of understanding the constitutive relationality of the self. The suitability of strong versions the second two types of constitutive relationality as models for constitutively relational autonomy has been rejected. That all of our ends should be

shared ends and not open to reconsideration is patently contrary what it means to be autonomous. Although we may ultimately accept many of the values we acquire from our societies, to be autonomous it must be at least possible to take a critical stance with the corresponding potential for rejecting these values (Barclay 2000, 63-4, Christman 2004, 154). That we should be defined by particular kinds of social relationships has been deemed either too potentially oppressive— again on the grounds of the impossibility of reconsidering our involvement in such relationships (Barclay 2000, 67-8) — or perfectionist (Christman 2004, 158).[86] So presumably, the relational nature of autonomy derives from how it is shaped (like our goals, beliefs, and values) by our social environment.

Westlund has offered an alternative account of the constitutive relationality of autonomy that relies on a relational conception of critical reflection rather than on finding oneself in the right kind of social environment (Westlund 2009, 35). On her view, which makes self-governance in the area of practical reason depend on one's openness "to engagement with the critical perspectives of others", autonomy is constitutively relational because the psychological capacities required for autonomy "point beyond" the agent (Westlund 2009, 35). While I have general doubts about whether Westlund's proposed "disposition to hold oneself answerable to others" does not reduce to more basic micro-level means, the more important criticism here is that this view represents a version of constitutive relationality in which no other actual people need participate since it can be satisfied through an engagement with imaginary others (Westlund 2009, 36).

---

[86] The variability here stems from differences in the *kinds* of relationships that are supposed to constitute us. Barclay has in mind communitarian arguments that claim that our roles as brothers, mothers, citizens of particular places are necessary parts of us. Christman has in mind arguments that Marina Oshana's claim that being autonomous means "standing in proper social relations".

Whereas Westlund locates the constitutive relationality of autonomy in defining critical rationality in a way that invokes the existence of other perspectives and therefore other people, these may only be potential or imaginary. My account represents a different way of understanding constitutive relationality. It does so by articulating how other people can, and often are, drawn into others' exercise of autonomy. Thus, in contrast to the view that social features *impact* the attainment and continued exercise of autonomy, my view is that on a practical level, social factors are very often *part* of our autonomy. Our exercise of autonomy is at least in part *constituted* by our relationships with others.

While my view falls short of Christman's standard for what would count as a "uniquely" (i.e. constitutively) relational conception of autonomy, this is at least partly because I resist the idea that necessary and sufficient conditions can be articulated for autonomy. Christman, conversely, seems to think that if autonomy is constitutively relational, relationality must feature among the "defining conditions" of autonomy as "conceptually necessary requirements for autonomy" (Christman 2004, 147-8). While I cannot show that relationality is conceptually necessary in this way, I have illustrated how, for many people, relationships with others will be a necessary part of their exercise autonomy in at least some contexts. And, if we take seriously Eva Kittay's remark that independence is largely a fiction "regardless of our abilities or disabilities" (Kittay 2001, 570), then I would hazard that potentially enough of us rely on others in our exercise of autonomy in enough contexts to go a significant way to approaching the level of generality that Christman thinks is desirable.

Third, I would like to highlight what I take to be two implications of my view for our judgements about the justifiability of paternalistic social responses to those who are found to have significant autonomy deficits. The first implication of my view is that the existence of deficits relating to the

personal capacities that are involved in the exercise of autonomy cannot be immediately taken to entail that a person is not autonomous. With automaticity, substitution, and reinforcement included as part of my account of the micro-level of analysis, we see that it is possible for deficits to be at least partly neutralized so that their effect on a person's functioning as it relates to autonomy may remain perfectly adequate despite their limitations. In addition, even setting aside automaticity, substitution, and reinforcement, because I have taken seriously the idea that people come to qualify as autonomous at different thresholds depending on the circumstances, people with deficits relating to their ability to exercise autonomy will not necessarily fail to pass the threshold for autonomy in all instances. The limits someone experiences in their ability to exercise the functions involved in autonomy may be more or less relevant depending on the context. If they can perform above the threshold for autonomy in a particular situation despite their impairments, then the fact such an impairment exists is of less relevance than in situations where they need to overcome the impairment to surpass the threshold for autonomy.

The second implication my view for paternalistic social responses stems from my insistence on paying attention to the full spectrum of autonomy functioning. In doing this it will be apparent that even though someone may not meet the threshold of autonomy required in a particular situation, they likely still exercise at least some level of autonomy functioning. It is my contention that respect for autonomy requires that this functioning, even though it is below the autonomy threshold in a particular situation, needs to be accounted for in our social responses. In general[87], people should be provided the place to exercise autonomy to the extent they are able, even if they are not capable of, for example, understanding the full complexity of their situation.

---

[87] I can envision some cases where this will not be acceptable. For example, concerns about public safety may over-rule preserving the space for someone with Alzheimer's disease, who tends to be forgetful and become confused, to continue driving "to the extent they are able". It may not be possible to enforce socially supported driving for this person, and the deficits they face may cause them to endanger others.

This aligns with calls among those who study the experiences of people with intellectual disabilities to recognize that autonomy may be exercised through deciding how much one wants to be involved in decision making that exceeds one's capacity, or deciding who should make decisions one's behalf (Lotan and Ells 2010, 116). It also provides further theoretical grounding for social interventions that create conditions in which people are supported in exercising the limited autonomy they do have, as opposed to making an immediate jump from non-intervention to coercive paternalism (Lotan and Ells 2010, 114, McLeod and Sherwin 2000, 268).

Overall, I see my view as providing the resources for avoiding the improper exclusion or underestimation of people as relates to their ability to exercise autonomy. Since my view does not conceive of the exercise of autonomy as something that must be achieved independently or perfectly, it provides the grounds for believing that more people are able to exercise autonomy than some other views might suggest. In so doing, I provide the grounds for further constraining the sphere in which paternalistic action might be justified. I do so in two ways: first, by showing that deficits in the skills and attitudes one needs in order to exercise autonomy do not automatically rule out autonomy; and second, by providing a conception of autonomy that is consistent with the idea that social arrangements should be such that they permit and support people in exercising the autonomy they do have, even if that autonomy is imperfect and insufficient for some contexts.

Bibliography

Anderson, Joel. 2014. "Regimes of Autonomy." *Ethical Theory & Moral Practice* 17: 355-368.

Arpaly, Nomy. 2002. *Unprincipled Virtue: An Inquiry Into Moral Agency.* New York: Oxford University Press.

Barclay, Linda. 2000. "Autonomy and the Social Self." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 52-71. New York: Oxford University Press.

Barkley, Russell A. 2000. *A New Look at ADHD: Inhibition, Time, and Self-Control.* New York: The Guilford Press.

Barkley, Russell A., Kevin R. Murphy, and Mariellen Fischer. 2008. *ADHD in Adults: What the Science Says.* New York: The Guilford Press.

Baumann, Holger. 2008. "Reconsidering Relational Autonomy: Personal Autonomy for Socially Embedded and Temporally Extended Selves." *Analyse & Kritik* 30: 445-468.

Beauchamp, Tom L. 2005. "Who Deserves Autonomy, and Whose Autonomy Deserves Respect." In *Personal Autonomy: New Essays on Personal Autonomy and its Role in Contemporary Moral Philosophy*, edited by James Stacey Taylor, 310-329. New York: Cambridge University Press.

Benson, Paul. 1994. "Free Agency and Self-Worth." *The Journal of Philosophy* 91 (12): 650-668.

Benson, Paul. 2005. "Taking Ownership: Authority and Voice in Autonomous Agency." In *Autonomy and the Challenges of Liberalism*, edited by John Christman and Joel Anderson, 101-126. New York: Cambridge University Press.

Berlin, Isaiah. 2002. "Two Concepts of Liberty." In *Liberty: Incorporating 'Four Essays on Liberty'*, edited by Henry Hardy, 167-227. New York: Oxford University Press.

Bermúdez, José Louis. 2005. *Philosophy of Psychology: A Contemporary Introduction.* New York: Routledge.

Biederman, Joseph, Michael C. Monuteaux, Alysa E. Doyle, Larry J. Seidman, Timothy E. Wilens, Frances Ferrero, Christie L. Morgan, and Stephen V. Faraone. 2004. "Impact of Executive Function Deficits and Attention-Deficit/Hyperactivity Disorder (ADHD) on Academic Outcomes in Children." *Journal of Consulting and Clinical Psychology* 72 (5): 757-766.

Blumenthal-Barby, J. S. 2016. "Biases and Heuristics in Decision Making and Their Impact on Autonomy." *The American Journal of Bioethics* 16 (5): 5-15.

Blumenthal-Barby, J. S., and H. Krieger. 2015. "Cognitive biases adn heuristics in medical decision making: a critical review using a systematic search strategy." *Medical decision*

making: an international journal of the Society for Medical Decision Making 35 (4): 539-57.

Calhoun, Cheshire. 2015. *Moral Aims: Essays on the Importance of Getting it Right and Practicing Morality with Others.* New York: Oxford University Press.

Christman, John. 1991. "Autonomy and Personal History." *Canadian Journal of Philosophy* 21 (1): 1-24.

Christman, John. 1989. "Introduction." In *The Inner Citadel: Essays on Individual Autonomy*, edited by John Christman, 3-23. New York: Oxford University Press.

Christman, John. 2004. "Relational Autonomy, Liberal Individualism, and the Social Constitution of the Self." *Philosophical Studies* 143-164.

Christman, John. 2014. "Relational Autonomy and the Social Dynamics of Paternalism." *Ethical Theory and Moral Practice* 17: 369-382.

Code , Lorraine. 2000. "The Perversion of Autonomy and the Subjection of Women." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 181-209. New York: Oxford University Press.

Colburn, Ben. 2011. "Autonomy and Adaptive Preferences." *Utilitas* 23 (1): 52-71.

Curatolo, Paolo, Romina Moavero, and Elisa D'Agati. 2010. "The Neurobiological Basis of ADHD." *Italian Journal of Pediatrics* 36 (79).

Dagger, Richard. 2005. "Autonomy, Domination, and the Republican Challenge to Liberalism." In *Autonomy and the Challenges to Liberalism*, edited by John Christman and Joel Anderson, 177-203. New York: Cambridge University Press.

Davy, Laura. 2015. "Philosophical Inclusive Design: Intellectual Disability and the Limits of Individual Autonomy in Moral and Political Theory." *Hypatia* 30 (1): 132-148.

Doris, John M. 2005. *Lack of Character: Personality and Moral Behaviour.* Cambridge: Cambridge University Press.

Doris, John M. 1998. "Persons, Situations, and Virtue Ethics." *Nous* 32 (4): 504-530.

Dubljevic, Veljko. 2013. "Autonomy in Neuroethics: Political and Not Metaphysical." *AJOB Neuroscience* 4 (4): 44-51.

Dworkin, Gerald. 1976. "Autonomy and Behaviour Control." *The Hastings Center Report* 6 (February): 23-28.

Dworkin, Gerald. 1981. "The Concept of Autonomy." In *Science and Ethics*, edited by Rudolph Haller. Rodopi Press.

Dworkin, Gerald. 1989. "The Concept of Autonomy." In *The Inner Citadel: Essays on Individual Autonomy*, edited by John Christman, 54-62. New York: Oxford University Press.

Earp, Brian D., and David Trafimow. 2015. "Replication, Falisification, and the Crisis of Confidence in Social Psychology." *Frontiers in Psychology* 6 (621).

Feinberg, Joel. 1989. "Autonomy." In *The Inner Citadel: Essays on Individual Autonomy*, 27-53. New York: Oxford University Press.

Fine, Kit. 2010. "Towards a Theory of Part." *The Journal of Philosophy* 107 (11): 559-589.

Foddy, Bennett, and Julian Savulescu. 2010a. "A Liberal Account of Addiction." *Philosophy, Psychiatry, & Psychology* 17 (1): 1-22.

Foddy, Bennett, and Julian Savulescu. 2010b. "Relating Addiction To Disease, Disability, Autonomy, and the Good Life." *Philosophy, Psychiatry, & Psychology* 17 (1): 35-42.

Frankfurt, Harry. 1971. "Freedom of the will and the concept of a person." *Journal of Philosophy* 68 (1): 5-20.

Frankfurt, Harry. 1988. "Identification and Wholeheartedness." In *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, edited by David Schoeman, 27-45. New York: Cambridge University Press.

Frey, R. G. 2005. "Autonomy, Diminished Life, and the Threshold for Use." In *Personal Autonomy: New Essays on Personal Autonomy and its Role in Contemporary Moral Philosophy*, edited by James Stacey Taylor, 330-345. New York: Cambridge University Press.

Fujita, Kentaro. 2011. "On Conceptualizing Self-Control as More Than the Effortful Inhibition of Impulses." *Personality and Social Psychology Review* 15 (4): 352-366.

Goldman, A. 2000. "Folk Psychology and Mental Concepts." *ProtoSociology* 14: 4-25.

Govier, Trudy. 1993. "Self-Trust, Autonomy, and Self-Esteem." *Hypatia* 8 (1): 99- 120.

Gracia, Diego. 2012. "The Many Faces of Autonomy." *Theoretical Medical Bioethics* 33: 57-64.

Guglielmo, Steve, Andrew E. Monroe, and Betram F. Malle. 2009. "At the Heart of Morality Lies Folk Psychology." *Inquiry* 52 (5): 449-466.

Harman, Gilbert. 2000. "The Nonexistence of Character Traits." *Proceedigns of the Aristotelian Society* 100 (1): 223-226.

Harman, Gilbert. 1999. "Virtue Ethics and the Fundamental Attribution Error." *Proceedings of the Aristotelian Society* 99: 315-331.

Heal, Jane. 2012. "Mental Disorder adn the Value(s) of 'Autonomy'." In *Autonomy and Mental Disorder*, edited by Lubomira Radoilska, 3-25. New York: Oxford University Press.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. "The Weirdest People in the World?" *Behavioural and Brain Sciences* 33: 61-135.

Jackson, Frank Cameron, and Philip Pettit. 1990. "In Defence of Folk Psychology." *Philosophical Studies* 59: 31-54.

Jaworska, Agnieszka. 2005. "Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value." *Philosophy and Public Affairs* 28 (2): 105-138.

John Doris and The Moral Psychology Research Group. 2010. *The Moral Psychology Handbook.* New York: Oxford University Press.

Khader, Serene J. 2011. *Adaptive Preferences and Women's Empowerment.* New York: Oxford University Press.

Killmister, Suzy. 2013. "Autonomy and the Problem of Socialization." *Social Theory and Practice* 39 (1): 95-119.

—. 2018. *Taking the Measure of Autonomy: A Four-Dimensional Theory of Self-Governance.* New York: Routledge.

Kingsbury, Justine, and Tim Dare. 2017. "The Philosophical Use and Misuse of Science." *Metaphilosophy* 48 (4): 449-466.

Kittay, Eva Feder. 2001. "When Caring is Just and Justice is Caring: Justice and Mental Retardation." *Public Culture* 13 (3): 557-579.

Knight, Amber. 2015. "Democratizing Disability: Achieving Inclusion (without Assimilation) through "Participatory Parity"." *Hypatia* 30 (1): 97-114.

Lefevor, G. Tyler, Blaine J. Fowers, Soyeon Ahn, Samantha F. Lang, and Laura M. Cohen. 2017. "To What Degree Do Situational Influences Explain Spontaneous Helping Behaviour? A Meta-analysis." *European Review of Social Psychology* 28 (1): 227-256.

Levy, Neil. 2006. "Autonomy and Addiction." *Canadian Journal of Philosophy* 36 (3): 427-447.

Lewis, David. 1972. "Psychophysical and Theoetical Identifications." *Australasian Journal of Philosophy* 249-258.

Lotan, Gurit, and Carolyn Ells. 2010. "Adults with Intellectual and Developmental Disabilities and Participation in Decision Making: Ethical Considerations for Professional-Client Practice." *Intellectual and Developmental Disabilities* 48 (2): 112-125.

Mackenzie, Catriona. 2000. "Imagining Oneself Otherwise." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 124-150. New York: Oxford University Press.

Mackenzie, Catriona. 2008. "Relational Autonomy, Normative Authority and Perfectionism." *Journal of Social Philosophy* 39 (4): 512-533.

Mackenzie, Catriona. 2014. "Three Dimensions of Autonomy." In *Autonomy, Oppression, and Gender*, edited by Andrea Veltman and Mark Piper, 16-42. New York: Oxford University Press.

Mackenzie, Catriona, and Natalie Stoljar. 2000. "Autonomy Refigured." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 3-31. New York: Oxford University Press.

Malle, B. F. 2006. "Of Windmills and Strawmen: Folk Assumptions of Mind and Action." In *Does Consciousness Cause Behaviour? An Investigation of the Nature of Volition*, edited by S. Pockett, W. P. Banks and S. Gallagher, 207-231. Cambridge, MA: MIT Press.

Maxwell, Scott E., Michael Y. Lau, and George S. Howard. 2015. "Is Psychology Suffering From a Replication Crisis?: What Does "Failure to Replicate" Really Mean?" *American Psychologist* 70 (6): 487-498.

McLeod, Carolyn, and Susan Sherwin. 2000. "Relational Autonomy, Self-Trust, and Health Care for Pations Who Are Oppressed." In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar. Oxford University Press.

Meyers, Diana Tietjens. 2005. "Decentralizing Autonomy: Five Faces of Selfhood." In *Autonomy and the Challenges of Liberalism*, edited by Joel Anderson and John Christman, 27-55. Cambridge, Mass.: Cambridge University Press.

Meyers, Diana Tietjens. 2000. "Intersectional Identity and the Authentic Self?: Opposites Attract!" In *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, edited by Catriona Mackenzie and Natalie Stoljar, 151-180. New York: Oxford University Press.

—. 1989. *Self, Society, and Personal Choice.* New York: Columbia University Press.

Milkman, Katherine L., Julia A. Minson, and Kevin G.M. Volpp. 2013. "Holding the Hunger Games Hostage at the Gym: An Evaluation of Temptation Bundling." *Management Science* 60 (2): 283-299.

Nahmias, Eddy. 2007. "Autonomous Agency and Social Psychology." In *Cartographies of the Mind: Philosophy and Psychology in Intersection*, edited by Massimo Marraffa, Mario De Caro and Francesco Ferretti, 169-185. Dordrecht: Springer.

Narayan, Uma. 2002. "Minds of Their Own: Choices, Autonomy, Cultural Practices, and Other Women." In *A Mind of One's Own: Feminist Essays on Reason and Objectivity*, edited by Louise M. Antony and Charlotte E. Witt, 418-432. Boulder: Westview Press.

Nussbaum, Martha C. 2011. *Creating Capabilities: The Human Development Approach.* Cambridge, Mass.: Belknap Press.

Nussbaum, Martha C. 1995. "Human Capabilities, Female Human Beings." In *Women, culture, and development: a study of human capabilities*, edited by Marth C. Nussbaum and Jonathan Glover, 61-104. Oxford: Clarendon Press.

Nussbaum, Martha. 2009. "The Capabilities of People with Cognitive Disabilities." *Metaphilosophy* 40 (3-4): 331-351.

Piper, Mark. 2016. "Achieving Autonomy." *Social Theory and Practice* 42 (2): 767-792.

Redley, Marcus, and Darin Weinberg. 2007. "Learning Disability and the Limits of Liberal Citizenship: Interactional Impediments to Political Empowerment." *Sociology of Health and Illness* 29 (5): 767-786.

Rips, Lance J., and Fred Conrad. 1989. "Folk Psychology of Mental Activities." *Psychological Review* 96 (2): 187-207.

Ryan, R. M., and Edward L. Deci. 2006. "Self-Regulation and the Problem of Human Autonomy: Does Psychology Need Choice, Self-Determination, and Will?" *Journal of Personality* 1557-1586.

Saigle, Victoria, Veljko Dubljevic, and Eric Racine. 2018. "The Impact of a Landmark Neuroscience Study on Free Will: A Qualitative Analysis of Articles Using Libet and Colleagues' Methods." *AJOB Neuroscience* 29-41.

Schwab, Abraham P. 2016. "Applying Heuristics and Biases More Broadly and Cautiously." *American Journal of Bioethics* 16 (5): 25-27.

Snow, Nancy E. 2013. ""May You Live in Interesting Times": Moral Philosophy and Empirical Psychology." *Journal of Moral Philosophy* 10 (3): 339-353.

Spellman, Barbara A. 2015. "A Short (Personal) Future History of Revolution 2.0." *Perspectives in Psychological Science* 10 (6): 886-899.

Stace, W. T. 1953. *Religion and the Modern Mind.* London: Macmillan.

Statistics Canada. 2017. *Highlight Tables- Age and Sex, Census year 2016.* May 3. Accessed January 31, 2018. http://www5.statcan.gc.ca/olc-cel/olc.action?objId=98-402-X2016002&objType=46&lang=en&limit=0.

Stich, Stephen P., and Shaun Nichols. 2003. "Folk Psychology." In *The Blackwell Guide to Philosophy of Mind*, edited by Stephen P. Stich and Ted A. Warfield, 235-255. Malden: Blackwell Publishing.

Stoljar, Natalie. 2000. "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.

Sumner, L Wayne. 1996. *Welfare, Happiness, and Ethics.* Toronto: Oxford University Press.

Westlund, Andrea C. 2009. "Rethinking Relational Autonomy." *Hypatia* 24 (4): 28-49.

Westlund, Andrea C. 2003. "Selflessness and Responsibility for Self: Is Deference Compatible with Autonomy?" *The Philosophical Review* 112 (4): 483-523.

Wolf, Susan. 1987. "Sanity and the metaphysics of responsibility." In *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, edited by Ferdinand David Schoeman, 46-62. Cambridge, Mass.: Cambridge University Press.

Wolf, Susan. 1989. "Sanity and the Metaphysics of Responsibility." In *The Inner Citadel: Essays on Individual Autonomy*, edited by John Christman, 137-151. Toronto: Oxford University Press.