



Kent Academic Repository

Porter, Elliot (2023) *Autonomy as an Ideal for Neuro-Atypical Agency: Lessons from Bipolar Disorder*. Doctor of Philosophy (PhD) thesis, University of Kent,.

Downloaded from

<https://kar.kent.ac.uk/100277/> The University of Kent's Academic Repository KAR

The version of record is available from

This document version

UNSPECIFIED

DOI for this version

Licence for this version

CC BY (Attribution)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal**, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

University of Kent

Autonomy as an Ideal for Neuro-Atypical Agency: Lessons from Bipolar Disorder

By Elliot Porter

A Dissertation submitted to the School of European Culture and Languages in
partial fulfilment of the requirements for the award of degree of Doctorate in
Philosophy

Canterbury, December 2022

Words: 96 005

Acknowledgements.

I am indebted to Lubomira Radoilska, for continuous support and guidance as my supervisor, and to Ben Colburn, whose mentorship and friendship began long before this project. I am also grateful to many colleagues who were members of the research network for the *Epistemic Injustice, Reasons, and Agency* project, and members of the *Philosophers with Lived Experience of Mental Illness* network. Both environments have been both nourishing and challenging, and allowed me invaluable space to develop this project, and to develop as a philosopher.

This work was facilitated by the Vice Chancellor's Scholarship, from the University of Kent.

Abstract

There is a strong presumption that mental disorder injures a person's autonomy, understood as a set of capacities and as an ideal condition of agency which is worth striving for. However, recent multidimensional approaches to autonomy have revealed a greater diversity in ways of being autonomous than has previously been appreciated. This presumption, then, risks wrongly dismissing variant, neuro-atypical sorts of autonomy as non-autonomy. This is both an epistemic error, which impairs our understanding of autonomy as a phenomenon, and a moral error, which withholds recognition and respect that autonomous agents are due. I argue that careful attention to the different 'shapes' of agent's autonomy reveals that there are indeed distinctive kinds of neuro-atypical autonomy that are widely mistaken for non-autonomy. This project argues for two propositions. First, that there is a kind of autonomy, with a distinctive shape and texture, available to bipolar agents. This is a variant sort of autonomy, but not a defective or deficient sort. It is the ideal of autonomy that it is worth bipolar agents striving for, given the range and intensity of experiences that they will have to confront. Whilst bipolar autonomy will look unlike more neurotypical kinds of autonomy, it continues to be owed recognition. The second proposition is that we should understand the ideal of autonomy to be both pluralist and highly context sensitive. Given the range of starting points from which agents will strive to an ideal of autonomy, there will be some variation in the ideals worth striving for. The variation in ideals will reflect the existential realities they are local to. Theory of autonomy must be adapted to recognise bipolar, and other neuro-atypical, kinds of autonomy, and to reflect this wider pluralism. These propositions will be argued for with careful attention to the experiences and agency of people living with bipolar, or related, disorders. Understanding the ways that manic or depressive episodes, and the shifts between them, alter and apply pressure to our agency will inform an account of the distinctively bipolar autonomy that can be built on top of it.

Table of Contents

Title Page	1
Acknowledgements	2
Abstract	3
Table of Contents	4
Introduction	7
<i>0.1 The Project</i>	<i>7</i>
<i>0.2 The Motivation</i>	<i>14</i>
<i>0.3 The Methodology</i>	<i>17</i>
<i>0.4 Chapter Plan</i>	<i>23</i>
Chapter 1 - Depression and Futility: Moral Psychology in a Normative Desert	27
<i>1.1 Introduction</i>	<i>27</i>
<i>1.2 Sentimentalism</i>	<i>28</i>
<i>1.3 Depression and Futility</i>	<i>31</i>
<i>1.4 Normative Outlooks</i>	<i>35</i>
<i>1.5 Regulative Guidance</i>	<i>42</i>
<i>1.6 Accidic Depression</i>	<i>47</i>
<i>1.7 What of Autonomy</i>	<i>50</i>
<i>1.8 Conclusion.</i>	<i>55</i>
Chapter 2 - Mania and Urgent Agency	57
<i>2.1 Introduction</i>	<i>57</i>
<i>2.2 Portrait of Mania</i>	<i>58</i>

<i>2.3 Manic Agency</i>	65
<i>2.4 Modular Agency and Agential Structures</i>	71
<i>2.5 Manic Agency as Urgent Agency</i>	73
<i>2.6 What Kind of Agency</i>	81
<i>2.7 Conclusion</i>	87
Chapter 3 – A Moral Epistemology for the Aftermath of Mania	89
<i>3.1 Introduction</i>	89
<i>3.2 The Aftermath Problem</i>	90
<i>3.3 The Positive Case</i>	99
<i>3.4 Mania Sensitises us to Value</i>	104
<i>3.5 Thick Fluency</i>	111
<i>3.6 Modes of Agency</i>	116
<i>3.7 Advantages of Manic Insight</i>	119
<i>3.8 Conclusion</i>	123
Chapter 4 – Self-Narrative and Self-Trust in Bipolar Disorder	126
<i>4.1 Introduction</i>	126
<i>4.2 Self-Trust</i>	127
<i>4.3 BoRD and Self-Trust</i>	132
<i>4.4 Adequate Narratives</i>	134
<i>4.5 Bipolar Narratives</i>	137
<i>4.6 The Medicine Journey</i>	143
<i>4.7 Securing Autonomy by Self-Narration</i>	148
<i>4.8 Conclusion</i>	151

Chapter 5 – Living Between Normative Worlds: Predicating Self-Determination	153
at Scope	153
<i>5.1 Introduction</i>	<i>153</i>
<i>5.2 Affective Shifts and Self-Determination</i>	<i>155</i>
<i>5.3 Two Self-Management Regimes</i>	<i>160</i>
<i>5.4 Wide Scope Self-Determination</i>	<i>165</i>
<i>5.5 Narrow Scope Self-Determination</i>	<i>169</i>
<i>5.6 What does this tell us?</i>	<i>174</i>
<i>5.7 Conclusion</i>	<i>179</i>
Chapter 6 - Acts, Descriptions, and Responsibility as Answerability in Mania	182
<i>6.1 Introduction</i>	<i>182</i>
<i>6.2 Marginal Responsibility</i>	<i>184</i>
<i>6.3 Answerability First</i>	<i>187</i>
<i>6.4 Fixing Act Descriptions</i>	<i>198</i>
<i>6.5 Constitutively Social Autonomy</i>	<i>211</i>
<i>6.6 Conclusion</i>	<i>217</i>
Conclusions	219
<i>7.1 Bipolar Autonomy is an Existent Variety of Autonomy</i>	<i>219</i>
<i>7.2 Recognising Bipolar Autonomy</i>	<i>224</i>
<i>7.3 Implications</i>	<i>227</i>
Bibliography	234

Introduction

§0.1 - The Project

This enquiry considers what autonomy is available to agents who live with bipolar or related disorders (hereafter, BoRD). We will examine the effects that BoRD has on someone's agential capacities and consider what this will do to their personal autonomy. The enquiry will motivate two propositions. The first is that bipolar autonomy exists as a distinctive kind of personal autonomy, available to agents living with BoRD. Bipolar autonomy is an ideal that is worth bipolar agents striving for, and can fill the same role that that general, (neuro)typical, or unspecified personal autonomy fills in moral and political theorising. Bipolar autonomy is distinct from other kinds of personal autonomy in that it responds to a distinctive *existential reality*, and in that it differs in its *shape and texture*. (Both phrases will be defined below). The second proposition is that we should revise existing theories of autonomy, using conceptual tools that will be developed during this enquiry, so that they will recognise bipolar autonomy.

§0.1.1 Bipolar or Related Disorders

BoRD is a medical term that picks out a cluster of related diagnoses, which in turn pick out clusters of symptoms that might be observed or reported in clinical contexts. They are "episodic mood disorders defined by the occurrence of Manic, Mixed or Hypomanic Episodes or symptoms. These typically alternate over the course of these disorders with Depressive Episodes or periods of depressive symptoms." (ICD-11, 2019, 6A6). These disorders are episodic in that depression, hypomania, and mania come in episodes that remit and relapse, (however we shall complicate this description in a moment). The emergence and remission of these episodes are the affective shifts that characterise BoRD. These affective shifts can have a sustained and disruptive effect on a person's life, projects, and relationships. We shall see this in some detail in Chapters 1 and 2, but it must be noted that a wholly episodic characterisation is misleading. Many people who live with BoRD experience symptoms of depression or mania outwith any discreet episode. An episode is a period of time when these symptoms, which might be ubiquitous, cluster and reach a particular severity. Symptoms of depression and mania might occur at the same time, and should they reach adequate severity, they are referred to as a mixed affective episode, (ICD-11, 2019, 6A60.9). This clinical language is useful for drawing a circle

around a set of phenomena that we are interested in, but we should be careful not to rely too heavily on the clinical conceptualisation. Being interested in autonomy, we will use the term BoRD to refer a little more expansively than the ICD-11 does. A diagnosis of schizoaffective disorder (ICD-11, 2019, 6A21) is not included in the same diagnostic class as BoRD, however is similarly marked by episodes of mania, hypomania, or depression. As bipolar and schizoaffective agents face similar obstacles to establishing robust autonomy, much of what is found in this enquiry will cross-apply. Our use of the term BoRD is not tied narrowly to the clinical concept but is used to pick out the *existential reality* that agents who live with these conditions have no choice but to face, but which might also be faced by agents who do not have or merit a diagnosis.

The affective shifts that mark out BoRD create obstacles for bipolar agents in establishing robust personal autonomy. Affective shifts disrupt one's sense of the value of things, alter motivation, and change one's sense of what is possible, likely, or worthwhile. This is probably undeniable, but on its own is not enough to seriously challenge personal autonomy. The gap between the bipolar and neurotypical agents' experiences in this regard is of degree and not kind. Yet the presumption that mad, mentally disordered, or specifically bipolar experiences endanger autonomy remains commonplace. 'Madness' occupies a space in the cultural imagination in opposition to 'reason', 'rationality', and 'legitimacy'. The identity 'mad person' is marginalising, particularly in that it strips mad people of the authority to interpret their own experiences, or contribute to collective enquiries, (Crichton et al 2017, Houlders et al 2021, Chapman & Carel, 2022). In the philosophical literature, this presumption is widespread. Madness or mental disorder routinely feature as negative examples used to mark out the boundaries of autonomy on this or that theory. Many theories are constructed in such a way as to issue in the conclusion that, whilst not *ipso facto* injurious to autonomy, the notable features of BoRD comprise defeating threats to autonomy.¹ Challenges to this presumption have not significantly arrested the practice of madness being used as the negative example, nor has a full articulation been offered of what is being missed due to this mistake. This is what our enquiry shall offer.

¹ Instances include, Edwards (1981), Feinberg (1989, 28), Van Willigenburg & Delaere (2005, compare Bolton & Banner 2012), Oshana (2006, 7), Beauchamp & Childress (2013, 105). Compare; Radoilska (2012a, x), Bortolotti (2013).

§0.1.2 Autonomy

For the purposes of this enquiry, we shall take autonomy to refer to a set of valuable capacities which are exercised and recognised, and to an ideal that is worth striving towards. The capacities are a set of agential capacities that are valuable, both to possess and to exercise, for the kinds of beings that have them. Agentive capacities are the reasons-responsive capacities which mean that some entities to count as agents, and want of which means other entities do not. Feinberg (1989) distinguishes these capacities from the *condition* of autonomy, (31). If one has the capacities required for autonomy but does not exercise them, or is prevented from exercising them, one does not enjoy the actual condition of autonomy. To reduce the number of things we will refer to as ‘autonomy’, let us say that possessing the relevant capacities, whatever they are, makes one a candidate for autonomy, and one is autonomous insofar as one enjoys the condition of autonomy. The condition of autonomy is at least partly contingent on one’s having and exercising these agential capacities. This is the sense in which autonomy is an agency concept. A debate continues as to whether (or to what extent) autonomy is constitutively social. Most scholars agree that some social conditions are *formatively* necessary for autonomy, without which we could not develop the relevant capacities, but another body of literature holds that the condition of autonomy can only exist if certain social relations hold between agents. At the far end of this wing, autonomy is solely a social status conferred on agents, and not a capacity that might be exercised or arrested at all. The majority of views fall somewhere on the wide spectrum in between pure capacity views and purely relational views. Like most scholars, I take autonomy to belong in-between the extremes, but I lean towards the relational side. Doubtless, autonomy involves a set of capacities, which are valuable in their exercise, however neither the enjoyment nor exercise of these capacities, even to a high degree, is sufficient for us to be satisfied that someone is living autonomously.

I take autonomy to be constitutively relational at least insofar as it requires that someone’s agential capacities are recognised in the correct way. In broad strokes, this is captured by the distinction Mitova (2020) draws between empowering and disempowering descriptions of actions. When answering the question *why* someone did something, we might recite the reasons that person took to speak in favour of what they did, or we might explain with reference to their mood, the fact that they had missed lunch, or a prejudicial social identity (the ‘erratic emotionality of woman’ for instance, 771). The former description recognises the person as enjoying an independent perspective on reasons and highlights their capacities to handle and respond to reasons as relevant to what happened. Such descriptions

ascribe power to them.² The latter description casts the person as the vehicle by which moods, neurochemistry, or women's *from-Venus-ness* act upon the world, conspicuously to the exclusion of the agent's influence. This is disempowering. Autonomy is constitutively social insofar as recognition of our agential capacities ascribes social and normative power to us. The uncontentious way this comes about is that others recognise us as having a point of view on practical reasons, and opting for descriptions of what we do, what would be good for us, and our contribution to social practices, in the empowering terms of our agency. We can also empower ourselves by using these descriptions, prompting others to see us as agents, or when we reflect on our own deeds, prompting self-recognition.

What this means is that the condition of autonomy requires not only that we have or exercise agential capacities, but that our possessing and exercising these capacities is recognised as agency at work. This is why the phrase 'autonomous agency' is not merely repetitive. There is a fact of the matter whether agency is at work, even if no one recognises it. This agent is a candidate for autonomy in virtue of these capacities but cannot be autonomous until they and their interlocutors recognise that agency for what it is. This recognition is constitutively social, and is a necessary condition for autonomy, but it is not sufficient. We might confer something like this recognition on entities that lack any such capacities. In his anthropological study of ethics, Laidlaw (2014) details the role played by responsibility attributions, as making a thing that happened to us into a thing that was done to us. Holding us responsible for other people's sickness, for example, and demanding restitution assigns us the role of helping that person through their illness.³ Agential capacities are no more present here than in the waters of the Dardanelles that swept away Xerxes' bridge. In either case, agency is *ascribed* to an entity (a member of the community, the Hellespont Strait) but is not *recognised* because the relevant capacities are not there to receive recognition. Humans cannot send their feelings out to harm people. Bodies of water are doggedly neutral in matters of geopolitics. Commanding that the Hellespont be whipped as punishment is appropriate given some situationally relevant norm (great kings must show that they are not impotent before nature, they can at least retaliate), but it does not make the waters autonomous.

² They answer specifically the Anscombian *why* question, (Anscombe, 2000, §5).

³ This holds true even when recrimination is inappropriate because we would not know that it is our unconscious resentment leaking out and attacking our neighbours. He notes that a standard part of the rite of healing is for the supposed witch to state that *if* it is their own feelings that are causing the sickness, they abjure them (207).

Autonomy requires both the presence and exercise of agential capacities, and social recognition of them. I do not take these to be jointly sufficient for autonomy. Vast literatures argue for more robust and specific conditions on either side: agential capacities must meet some standard or be exercised in some way, social arrangements must be such and so, before autonomy obtains. I take these to be the most basic conditions required for an individual to enjoy autonomy. Remaining as neutral as possible between different theories of autonomy is necessary if we are to succeed in our second objective, which is offering the tools required to revise a significant portion of existing theories. However, as we must have some model in place to proceed, we shall understand the condition of autonomy to obtain when agential capacities are present, exercised, and recognised, *mutatis mutandis* as specified by whichever is eventually your preferred theory.

Autonomy, as capacity-cum-value, refers to the presence, exercise, and recognition of those capacities that are valuable. However, autonomy is also an ideal worth striving towards. We rightly hope to enjoy the condition of autonomy, and for others to enjoy that condition. Autonomy admits of degrees, but as an ideal, it is the condition we are justified in striving for and in supporting others to strive for.

“Autonomy... is not merely a “condition,” but a condition to which we aspire as an ideal.” (Feinberg, 1989, 31). The first proposal that this enquiry supports is that bipolar autonomy exists as a distinct kind of autonomy from more familiar neurotypical kinds. It is in this sense, autonomy as an ideal worth striving for, that this proposal holds true. Bipolar agents remain candidates for autonomy, as we shall see, because BoRD alters but does not eliminate their agential capacities. It might be thought that this change leaves bipolar agents as candidates for autonomy, but unable to achieve the condition to a high degree. (Indeed, this might contribute to why we refer to BoRD as disordered, as suggested by Radden, 2013). Against this position, we shall investigate what the ideal form of bipolar agency looks like on its own terms. Centring mad, and in this case, bipolar, perspectives will allow us to see autonomy in a new light and offers a more pluralist way of understanding autonomy.

§0.1.3 Autonomy Pluralism

BoRD will act upon a person’s agential capacities in a number of ways, and we must know that these capacities persist in a state that makes the person still a candidate for autonomy. It might be objected that whilst a bipolar agent remains a candidate for autonomy, there is a ceiling as to how highly that

autonomy can develop, imposed by the changes BoRD effects to their agentic capacities. To speak of more or less highly developed autonomy is to refer to how close some agent's condition comes to the ideal of autonomy. If, however, an agent's capacities are such that they cannot be in the condition described by our ideal of autonomy, we should ask whether that is the relevant ideal or not.

We can apply a broadly eudaimonist framework to our model of autonomy. The ideal condition of autonomy involves agentic capacities operating in an excellent way, in excellent circumstances (such as enjoying recognition, having many options, etc). The ideal of autonomy is, broadly, to flourish *qua* agent: to be an excellent example of an agent. However, if we are to motivate our first proposition, this flourishing state must admit of more plurality than is popular in some eudaimonist ethics. The classical Aristotelian model considers what an excellent example of a man looks like: a political animal with rational capacities, who exercises these capacities in ways conducive to his living well with others. On the far side of this tradition, Philippa Foot takes the subject of flourishing to be an organism, an entity defined in part by a set of functional norms. To flourish is to be an excellent example of the kind of organism one is, and so a human flourishes insofar as she pursues goods and reasons in accordance with a normative structure that is confined by her nature as an organism, (2001). In both cases, excellence in the kind of thing that we are turns on the description favoured for the kind of thing that we are (rational political animal, organism).

If the ideal of autonomy can be expressed as being an excellent example of an agent, then the description deployed in our account of agency matters greatly. The bipolar agent might struggle to ever achieve the ideal under a conventional description, but as Velleman points out (1992) the conventional ideal is often a *square*, both unrealistic and undesirable. However, if our agentic capacities can be restructured in various ways, as I suggest, then we can excel as agents under different descriptions. The narrower ideal of autonomy is to flourish *qua* agent, meaning to be an excellent example of the kind of agent we are. A key part of the project here is to specify what kind of agency is available to people living with BoRD, and so to tell us what kind of autonomy is worth striving for. This is not an ideal of autonomy that everyone should pursue. Rather, it reveals the differences between the ideal of autonomy in the face of the different existential realities faced by bipolar and neurotypical agents.

This approach to flourishing runs contrary to the typical Aristotelian move. Aristotle considers the features that make a harpist a good harpist or carpenter a good carpenter to be constitutive of the concept of 'harpist' or 'carpenter'. To know what makes a good human being then requires that we investigate what is constitutive of a human being, that rational political animal. Our status as humans is universal over harpists and carpenters, and so is taken to be fundamental in a way that gives us the authoritative picture of how we flourish. Making a similar move, Foot takes the description of us as organisms to be fundamental, and so to provide the more authoritative picture of flourishing. This is the move we avoid. Whilst the descriptions of 'organism' and 'human' are inescapable, we needn't rely on them to give us the most authoritative definition of flourishing. These descriptions are drastically underspecified, and so provide only broad constraints on what could be called flourishing for anything describable as a human organism. In the same way, 'agent' is too under-specific to tell us what ideal condition is worth striving towards given the kind of agent that we are. 'Organism' and 'human' are broad categories which inform what flourishing will look like, but which require more specification to give us an account of what this or that human person's flourishing consists in. The poet who is lauded by their community, taught in schools and universities, and who influences their literary community flourishes both as a human and as a poet. The archer who takes medals at the Olympic Games, releases their sports calendar for a few years in a row, and coaches the next Commonwealth youth team flourishes both as a human and as an archer. The poet would still flourish as a human were they to enjoy the archer's particular successes, but they would not be flourishing as a poet. The success she enjoys is a flourishing life for some human or other, but not for her. The purpose of this project is to better specify the agency at work in people who experience depression, mania, and BoRD, so that we can better understand the ideal of autonomy worth striving for, from this starting place. A narrower specification of what bipolar agency looks like will provide for us a clearer ideal of autonomy to strive for, that does not simply reflect an ideal of neurotypical agency.

I have used the term 'existential reality' to roughly divide up kinds of autonomy, and to carve out the corner we are interested in. One's existential reality is, in broad strokes, the set and range of experiences, particularly experiences of value, that one will have, can predict having, and has no choice but to have. Some range of experiences is an *existential* reality because it is an unavoidable feature of existing as the particular, or the kind of, being that one is. We might imagine that Velleman's *squares* enjoy a placid existential reality, involving experiences of desire unmixed with ambivalence or

ambiguity. Compare this with a predictable experience of outright contradiction in one's sense of value, where entire families of reasons will become invisible to one for a spell, and other families of (or individual) reasons will consume one's entire attention at other times. People with BoRD can expect to lose most or any sense of the value of things, for weeks or months at a time, during which time they will sincerely recant the value they placed on their long-standing projects. At other times, they will become so flighty, or fixated, that they will not be able to sustain the wide range of projects that make up most people's lives. This is existential as it speaks to how one interacts with the normative world and carves out one's place in it, and is someone's reality in that whatever theoretical perspective we might take on agency, and no matter how effective medical interventions and self-care prove to be, this is the set of experiences that will obtain. These existential realities must be confronted, and a neurotypical ideal of agency will frequently be inadequate to confront the existential reality of BoRD. Existential realities introduce non-optional descriptions that will alter what our flourishing looks like, and will change what the relevant ideal of autonomy is for agents in our shoes.

§0.2 - The Motivation

Bipolar agents confront an existential reality that is made distinctive by the repeated affective shifts they will encounter, which will drastically alter their sense of what is valuable, and create significant obstacles for their establishing conventional sorts of autonomy. To conclude that bipolar agents have no hope of reaching the ideal of autonomy is a counsel of despair, criticisable more for being unuseful than for being false. If bipolar agents have no hope of reaching a conventional or familiar ideal of agency, that condition is not what fills the 'to be striven towards' role for those agents. This role is what makes some condition the ideal relevant to its domain. To ask what autonomy as an ideal involves is to ask what condition of agency we should strive towards. To hear the answer that bipolar agents cannot reach an ideal is to hear that bipolar agents should not bother striving for any particular condition of agency. We should respond to this the way we should respond to all counsels of despair: 'and what if *giving up* doesn't work?'

Bipolar agents still face the task of confronting their existential reality, of developing whatever agential capacities they have, and establishing a suitable agential project of their own. It seems unlikely that bipolar agents, uniquely, have completely free reign to do this unguided by anything worth striving towards. Rather, as with all other agents, there will be more and less valuable ways for bipolar agents to establish what autonomy is available to them. Given their existential reality, and the consequences that has for their agency, what bipolar agents have best reasons to strive towards will be different from that which neurotypical agents should strive for. Our ideal of autonomy, in other words, is local to the kinds of agency that are readily available to us. The negative conclusion, that ideal autonomy is unavailable to bipolar agents, evades the actual question. This project is motivated by a rejection of the negative conclusion and attention to the question that bipolar agents continue to face: how to develop agency that is compatible with, and which confronts, their existential reality. What is the ideal of autonomy local to their situation?

This enquiry has roots in wider neurodiversity and mad activist literatures, which highlight the problem with understanding concepts like 'agency' or 'autonomy' against an ideal set only by (neuro)typical expression. It is that marginal agents, whose experiences are reflected upon to better inform theory at the centre, may reasonably demand to know why they are at the margins. Indeed, marginal agents might build a detailed model of their own moral psychology, declare it to be ideal, and declare neurotypical agents to be marginal: tragically hamstrung by their dulled appreciation of the importance of the things that manic or autistic people fixate on, and impaired by their inability to hyperfocus. If a slogan was needed to capture the focus of mad activism, Jim Sinclair's entreaty would serve amply: "[g]rant me the dignity of meeting me on my own terms – recognize that we are equally alien to each other, that my ways of being are not merely damaged versions of yours" (1992, 302). Various theories of agency have enquired deeply, delving to the tectonics, and explaining typical or familiar kinds of agency to the most fundamental level. But agency is a warren that must be mapped. The typical or ideal agent we imagine in classic philosophy of action is local to only one part of that warren, and is frequently neither a desirable ideal nor a realistic neighbour. Velleman's *squares* are fringe agents in their own way, so far from the agency we would recognise in ourselves or those around us.

Two concerns are woven tightly together here. One is epistemic and one is moral. The epistemic concern is to do with the nature of autonomy in general, the degree of pluralism it admits of, and what it looks like in the context of BoRD. The moral concern is what Sinclair addresses – a question of dignity, and what is owed to bipolar, neurodivergent, and other mad people. Even when we untangle these concerns, we will discover that they cannot be wholly separated. As valuable as work in philosophy of action that delves to the tectonics has been, the presumption that mad sorts of agency are marginal or best deployed as negative examples has focused attention on a relatively narrow patch of the vast warren that needs to be mapped. If due attention is not paid to agency as it exists in mad spaces, our understanding of agency is needlessly limited, and our understanding of what kinds of agency could be worth striving for (which ideals of autonomy are on offer) are similarly restricted. This concern becomes morally significant when, for want of attention to these spaces, we do not understand what ideal of autonomy it is worth bipolar agents striving for. Bipolar agents are then deprived of the understanding they need in order to choose to establish autonomy of some particular shape, and their friends, families, and clinicians lack the understanding required to support them in establishing and protecting that autonomy.

The absence of this understanding risks creating a *hermeneutic injustice*, whereby a community fails to produce the interpretive tools that a subgroup of that community needs in order to understand their own situation, due to the marginalisation of that subgroup (Fricker 2007, 152). Hermeneutic injustice is a species of broader *epistemic injustice*, the wrong that can be done to someone “in their capacity as a knower” (1). This concept has been expanded, and in recent years it has been suggested that different kinds of epistemic injustices are united by an underlying presumption:

[U]nprivileged knowers’ reasons for action and belief are treated as being ‘of the wrong kind’ by the default. In other words, their reasons are conceived as explanatory at best, but not as potential justifications. The upshot is to undermine agency overall since, in an important sense, unprivileged knowers are denied knowledge of what they are doing. Instead, privileged observers or critics are assumed to have better understanding of what such knowers are up to.
(EIRA 2018).⁴

⁴ The Epistemic Injustice, Reasons, and Agency project (2018-2022) was a collaborative research project co-lead by Professors Lubomira Radoilska (University of Kent) and Veli Mitova (University of Johannesburg). The cited passage is the central hypothesis of the project.

This is the connection that makes the epistemic and moral threads not wholly separable. Hermeneutic injustices occur when a subgroup's hermeneutic needs are not taken to be sufficiently reason-giving by the wider epistemic community, and so interpretive tools adequate for their experiences are not developed. The protests of these groups, and their efforts to generate these interpretive tools alone, are ignored, written off, or pathologized. Members of the marginalised group are not taken to be epistemic agents, capable of trading in and issuing justifying reasons. The disempowering presumption is not restricted to the epistemic domain, however. The tendency to pathologize the utterances of mad people is rooted in the same assumption of incompetence and illegitimacy that means mad people's decisions and actions are pathologized – revealing explanatory, but not motivating or justifying reasons. This is why we noted Mitova's framework earlier, and is where Sinclair's concern for dignity rests. We do agents and injustice when we seek to explain their behaviour in contexts when we should ask the Anscombian *why* question. In contrast, we respect people's dignity as agents when we seek out their view of reasons. Recognising that there are motivating reasons present acknowledges someone as exercising some agential capacities. For this reason, the concern of epistemic injustice is also a concern for the autonomy of bipolar agents.

§0.3 - The Methodology

This inquiry is a work in analytic philosophy, however as we must develop some new conceptual tools in order to understand our target phenomenon, the project is informed by a number of theoretic frameworks, some of which are a step removed from the analytic tradition. This enquiry will pay close attention to the first-hand phenomenology of depression, mania, and the shifts between them, and assumes that bipolar agents enjoy some privileged epistemic access to the kinds of agency they have at their disposal, and the kinds of autonomy that might be built upon it. The enquiry seeks to motivate a quite revisionary approach to understanding autonomy, according to which bipolar autonomy exists as a variant kind of autonomy, rather than sometimes existing as conventional autonomy *surviving* the condition. This revisionary tendency defers, in many cases, to the first-hand phenomenology of bipolar agents. This deference is not total. It is inevitable that sometimes bipolar agents will not have knowledge as to what they are doing, or transparently appreciate their own motivating reasons. As

much as possible we shall avoid revising bipolar agents' first-hand reported phenomenology. This does not mean that there is no space for interpretation to be done, but only that we dispense with the assumption, noted in the EIRA hypothesis, that bipolar agents lack knowledge of what they are doing. Let us note the theoretic frameworks that are deployed to offer methodological support to this enquiry.

§0.3.1 - Multidimensional Autonomy

Rather than concluding that BoRD simply limits or destroys autonomy - a council of despair for people who nonetheless must face the world - I will examine the abnormal shape of autonomy as it obtains in these contexts. Recent literature has come to understand autonomy as a multidimensional phenomenon (Mackenzie 2014, Killmister 2017). Autonomy can extend more or less far in a number of dimensions. The consequence of this is that autonomy will look very different between two people, extending as it does to drastically different extents in different dimensions. Suzy Killmister is the most explicit in making this her project. She is engaged, she tells us: "to shift the discussion from whether someone is autonomous, to the ways in which she is autonomous." (2017, 18).

How are we to compare two agents, one who has constituted a rich personal identity but who struggles with self-unification, and another with a truncated personal identity but who achieves a high level of self-realisation? One thing we can't do is say definitively who is the more autonomous. What I am offering here is not a score-keeping mechanism, which we can apply to get a precise readout of how autonomous a particular agent or action is. Instead, my hope is to offer a more complex theory which we can use to understand the precise shape of someone's autonomy

(Ibid, 17-8)

Killmister offers a topographical program that lets us map the shape and contours of an agent's autonomy. She is concerned with identifying the failures that impair someone's autonomy, as supporting a person's autonomy requires very different interactions when one is excessively deferential from when someone is physically coerced. The most useful work her program achieves, for our purposes, is mapping the shape of agencies that are mistaken to be non-autonomous, too marginal, or too fragile to do the normative work that autonomy takes up. Recognising that autonomy comes in many shapes and textures, however, will allow us to recognise expressions of autonomy that are far lumpier than we are used to, and so are likely to mistake for non-autonomy. Mapping the contours of different kinds of autonomy will make it easier to recognise variant kinds as autonomous, rather than mistaking them for defective (or non-) autonomy.

The shape of an agent's autonomy refers to the degree to which it extends in various dimensions. Different theories specify different numbers of dimensions, and so will differ on how they understand the shape of some agent's autonomy. We could think of each divot in autonomy, or dimension in which it does not extend as far as it might have, as a partial failing of autonomy, however we need not think this way. If autonomy is unavoidably lumpy, then these are features rather than failures in an agent's autonomy. Discovering the shape of an agent's autonomy, how far it extends in this or that dimension, need not evaluate their autonomy overall, only describe it. This description can be valuable in informing the agent of where they are and are not autonomous, where their autonomy requires support, and of what their starting position is should they chose to reshape their autonomy. It might be reasonable for an agent to give up some autonomy in one dimension in order to secure further extension in another. Indeed, choosing to craft one's autonomy into some particular shape would seem to be a deep and important exercise in self-authorship, self-creation, or however else we prefer to conceive of autonomy.

The texture of an agent's autonomy is a somewhat looser term which I offer to supplement Killmister's framework. Whatever are the dimensions of autonomy on your preferred theory, differences in shape are not the only differences that are worth knowing about. I use the term texture to refer to the qualitative differences that will hold between two agents' autonomy, which we might be interested in for the purposes of distinguishing bipolar from neurotypical kinds of autonomy. Some of the differences we note in bipolar autonomy will be specific to shape, how far autonomy extends in the dimensions used by whichever theory. However, many differences will not change the extension in some dimension but will remain interesting for our purposes. These are the textural differences that it would behove us to understand. The same feature of bipolar autonomy might be understood as a difference in shape one theory, and a difference in texture on another, depending on what are the dimensions of autonomy used by that theory. Whichever column that feature lands in, it will still contribute to a rich and detailed story about an individual's autonomy.

§0.3.2 - Epistemic Justice

In §0.2 we noted the role that concerns of epistemic justice played in motivating this enquiry. As a methodological framework, however, this point bears elaboration. As developed by the EIRA project,

epistemic injustices are underpinned by the assumption that some person or group of people cannot trade in justifying reasons, and so their testimony fails to be reason-giving with respect to others' beliefs, and their own accounts of their actions fail to reveal the reasons for which the actions were done. The most direct expression of this injustice is in the adopting of an explanatory stance towards someone, which seeks explanations for what they have done, where an agentic stance would be appropriate. As a methodological conceit, we will adopt an agentic, rather than explanatory stance towards the testimony of depressed or manic people writing about their own experiences. We shall ascribe knowledge of what was done, what reasons it was done for, and the context it was done in, to these agents, unless there is a compelling reason not to.

The assumption made here is that bipolar agents enjoy privileged epistemic access to knowledge about the reasons they respond to, the kinds of agency that are available to them, and what autonomy might look like in their contexts. This is the kind of knowledge agents have first hand, even if it requires a degree of self-interpretation to acquire. It is also knowledge we can struggle to acquire because we lack insight, feel compelled to preserve a certain self-image, or are discouraged from introspection by our peers. Bipolar agents will face these obstacles to self-knowledge in the same way that all agents do. However, where this self-knowledge is achieved, we can rely on the bipolar agent to share this knowledge. Understanding how the bipolar agent tracks and responds to reasons, how their agency is structured and restructured, and what ideal of agency they aspire towards, will inform our enquiry into the nature of bipolar autonomy. To gain access to this privileged knowledge requires, therefore, that we pay close attention to the first-hand phenomenology shared by bipolar agents. For this reason, our enquiry shall include detailed analyses of autobiographical writings about depression and mania. These writings are not usually offered for the purposes of informing philosophical theory of action, and so some interpretive work will be required to resolve what is said into a picture of how the depressed or manic agent experiences value, encounters reasons, and responds to them. This will allow us to track in fine details the distinctions that mark out agency in contexts of mania and depression.

To complement the phenomenological thread in this project we adopt a neo-Aristotelian framework that emphasises the role emotions play in rational motivation. BoRD, as an affective disorder, is marked out in particular by changes in the kinds of emotional experiences we are prone to have, and so

accounting for the role of these emotions in agency calls for a framework that recognises them as motivationally fertile, reasons-responsive states. We shall elaborate more on this framework in §1.2.

Whilst I suggest that BoRD creates an epistemically privileged standpoint, from which bipolar agents have an advantage in understanding variant kinds of autonomy, it is not a standpoint in the way more usually considered in standpoint epistemology. Conventionally, the standpoint is a social position from which one has privileged, if not exclusive, access to situated knowledge, (see Harding 2004 for an overview). This is knowledge about social relations that goes beyond the simple qualia of occupying that social position, but which is hard or impossible to access for those who share one's social order but not one's position in that order. Certainly, the identity 'mad person' is marginalising in precisely the way that gives rise to epistemically privileged standpoints, but this is not the standpoint that I suggest. It is not in virtue of bipolar agents' social positions that I take them to enjoy privileged epistemic access, but in virtue of being the agents with this kind of agency, who respond to reasons in this way, and whose autonomy will be shaped in these ways.⁵ Thus, I adopt a more general standpoint epistemology, according to which the bipolar agent has privileged access to knowledge about autonomy and agency in virtue of being situated in their particular kind of agency, complete with the challenges of that existential reality, and the insights provided by those experiences of value and reasons.

§0.3.3 - Medical Terms and Medical Conceptions

I have adopted clinical terms, but am concerned with agency, reasons, and the wide range of ways autonomy might be structured. It might reasonably be thought that clinical terms are unhelpful, due to their inbuilt tendency to pathologize. Nosological categories serve clinical and (psychological or neurological) research purposes to the extent clinicians and these researchers deem necessary. They are neither intended to nor capable of mapping comprehensively onto differences in our existential realities. I adopt this language because it is a useful, but far from perfect, way of bringing together the grab-bag of experiences that have some bearing on how a person flourishes, and which make this corner of the warren of agency particularly interesting. Diagnoses are sought out, or imposed, usually in

⁵ This is also disanalogous to those standpoint epistemologies which focus on gendered cognitive styles which emerge due to social position (eg, Hartsock, 1987), as it is not the style with which one pursues self-knowledge that grants greater access, but that it is one's own cognition that one is seeking to learn about.

circumstances when a person is struggling to flourish, has sought out support, or needs to name the thing that causes their distress. In some cases, this involves adopting the medical conception of BoRD, perhaps reifying the diagnosis into another agency to which thoughts feelings or actions might be attributed, and against which one must constantly guard some deep self. But in other cases, diagnoses are only banners one might fly to attract people with similar experiences in order to build community. Diagnoses can be the germ of a social identity that demands recognition.⁶ In some cases, it is adopted for the sake of having any language at all with which to discuss shared experiences.

It is in the latter register that I use these terms. To adopt the language of despair, mystic inspiration, and the multitudes which each of us contains, is to cast the net too widely, drawing in experiences that are nearly universal, but arise incidentally. Most people will experience each of these things throughout their lives, but do not live with the knowledge that they are on the horizon, and will need to be carefully managed so that they don't take over when they arrive. This is a difference in existential reality. Clinical language is a uniting point for mad people, and a useful standpoint from which to interrogate marginal status: 'are we at the margins of what constitutes agency, or what is recognised as agency?' This is our starting point for centring mad sorts of agency, which call for dignity and recognition on their own terms.

§0.3.4 - Theoretic Neutrality

As far as possible, I will remain neutral between specific theories and broader conceptions of autonomy. Theory-neutrality is necessary if the conclusions of this enquiry are to be useful across a wide range of theories. Were we to adopt some particular theory or conception of autonomy for the purposes of this enquiry, our conclusions would likely apply only over that theory and perhaps its near neighbours. Therefore, keeping commitments minimal will ensure our conclusions are more widely informative. What we can learn about autonomy generally, by paying careful attention to bipolar autonomy, will not be conditional on subscribing to a restrictive set of theoretic commitments. I make an exception in Chapter 5 for the concept of self-determination, which I take to be a dimension of autonomy for the purposes of our model. This is necessary as without adopting some model, we would be limited in how

⁶ See Rashed's excellent (2019) investigation on this point.

far we could examine the shaping and reshaping of autonomy, at the level of dimensional extension, that affective shifts occasion.

However, it is impossible to approach these questions with nothing in place, which is why I have adopted the minimal model outlined in §0.1.2. Wherever we look, autonomy will be an agency concept, and so some agential capacities must be present for autonomy to be at stake. I have adopted a constitutively social model of autonomy, in part to distinguish agency from autonomous agency, but also to ensure that our findings can be applied to that side of the theoretic spectrum. Theorists who do not subscribe to a constitutively social view of autonomy can disregard these features of our enquiry, and focus on the structure of depressive, manic, and broader bipolar agency that is revealed. This theoretic minimalism is required in order to ensure that our conclusions are widely applicable. Without some minimal model, the enquiry could not proceed, and without appropriate slots for different families of theories to plug into this model, the conclusions would be of restricted importance.

§0.4 - Chapter Plan

Chapter 1 will consider a familiar type of nihilistic depression, which is well attested in the autobiographical and phenomenological literature. This depression imposes a view of the world as normatively barren, lacking in value, and devoid of adequately good reasons to do anything or feel any way. This *futility thinking* has a pervasive effect on the workings of agency. Careful attention to the phenomenological reports, placed in a sentimentalist framework, will produce two resources that will be useful moving forwards. The first is an account of how agency operates when an agent is depressed in this way. We must understand this if we are to understand what agency might be the foundation for personal autonomy during periods of depression. Changes in this agential foundation will transmit up to any autonomy established upon it, and have consequences for the shape and texture of that autonomy. Attention to what changes between periods of remission and *futility thinking* will outline for us a general model of agency as reasons-responsiveness that we will be able to rely on going forwards. This chapter

will give us the first look at the appreciative mode of agency, which is a conceptual resource that will be developed in later chapters.

Chapter 2 will consider manic moods, which have received significantly less attention than depressive moods have in analytic philosophy. Applying the model developed in Chapter 1, this chapter will examine how mania restructures agency, with particular attention to the kinds of response that practical reasons seem to call out for. Understanding mania in terms of the pervasive urgency that, from the manic point of view, seems to attach to all worthwhile considerations, will fill in another side of our picture of the agency that underpins bipolar autonomy. This chapter will demonstrate clearly that whilst there is a difference in the kind of agency at work, it remains recognisably agency, understood in terms of reasons-responsiveness, and so invites understanding as variation, rather than defect. In conjunction with Chapter 1, this will satisfy the initial conditions for our first objective: to show that there exists a variant sort of bipolar autonomy. The initial conditions are that there are agential capacities at work and that they might enjoy recognition. Having illustrated these capacities and their operation in the immediate sense, we will be able to turn our attention to the conditions for diachronic agency.

Chapter 3 considers a problem that might emerge for people whose agency is sometimes restructured along the lines of manic urgency. They might become alienated from their own normative outlook at different times, either losing the self-trust that is necessary for autonomy, at least in its exercise, or fragmenting their agential life, surrendering a large chunk to 'mania', reified as some other entity that soaks up their own agency. This chapter argues for a particular approach to moral epistemology in the aftermath of manic episodes which addresses these problems, and allows the bipolar agent to deploy their manic agency as a constructive element in the kind of effective diachronic agency that can be the foundations for robust personal autonomy. This chapter will further develop our understanding of the appreciative mode of agency, and argue that agency restructured by mania is in fact very well specialised to this mode.

Chapter 4 will consider the threat affective shifts pose to autonomy in another respect. These shifts can injure the self-trust that is necessary for agents to enjoy autonomy on various theories, but which is also

necessary if agents are to exercise whatever agency they have in a way that provides suitable foundations for personal autonomy. This chapter will examine the self-narration practices that are adequate to address this problem and which permit bipolar agents to thread episodes of their agency together into the sort of diachronic agency that can be the foundations for robust personal autonomy. The narratives that are best suited to this task will include certain features that are likely to have some consequences for how agents relate to themselves and craft their autonomy. Whether or not a theory includes self-narrative as a feature or dimension of autonomy, different approaches to self-narration will have different consequences for the shape and texture of an agent's autonomy.

Having addressed challenges in establishing the kind of diachronic agency that might be the foundation for personal autonomy, Chapter 5 will consider consequences this kind of diachronic agency will have for the shape and texture of autonomy, focusing on self-determination as a candidate dimension, as it or something like it is popular as a dimension of autonomy. This chapter will demonstrate the tight relationship that holds between self-determination, and other likely dimensions of autonomy, and so highlight the set of choices and trade-offs that must be made if an agent is to actively shape their autonomy. This chapter will also highlight a distinction in the scope at which self-determination might be exercised which will lead to textural differences in autonomy that bipolar agents must make. Vocabulary that can track these finer-grain distinctions will allow us to discuss the features of bipolar autonomy that make it distinctive from more neurotypical kinds.

With a robust picture in place as to what agential capacities are present, how they thread together into an effective diachronic agency, and what kinds of support are suitable to keep them effective, we must consider how the social conditions of autonomy might be satisfied for bipolar agents. Chapter 6 offers an account of responsibility in mania that serves two purposes. The first is to demonstrate the practical applicability of what has gone before. Understanding manic agency as restructured and pervasively urgent has consequences for how relationships with manic agents can and should be conducted, and recognising the different scopes at which autonomy might be exercised makes it important for agents to shift perspective and look at what they and others have done at another resolution, sometimes, in order to understand the actions that have taken place. Independently of this, however, this account offers a practical avenue for bipolar agents to receive the social recognition of their agential capacities that is

necessary for the exercises of these capacities to become autonomous on the model we are using. A practical avenue is valuable as it does not rely on engagement with academic philosophical work before the peers and clinicians of bipolar agents confer recognition on the agentic capacities that are at work. Such an avenue to recognition can more widely satisfy the basic social conditions of autonomy, and so support the autonomy of a great many more bipolar agents.

Chapter 1 – Depression and Futility: Moral Psychology in a Normative Desert.

§1.1 Introduction

Depression is widely taken to threaten personal autonomy, particularly by limiting the scope of action. When depressed, we do little and give up on much. However autonomous we might have been, we will exercise that autonomy less frequently. Due to the changes in our apparent values, we might even enjoy less autonomy than we did.

This chapter develops an existing sentimentalist account of moral psychology with a view to making this kind of depressive inertia intelligible in response to the Anscombian *why* question, and considers the effects this depressive mode of agency will have on our autonomy. This account will establish the framework for discussing agency we will use during the remainder of this investigation. I focus on a particular sort of nihilistic depression as the key example of where depression impairs our agency, by causing us to fail to recognise many practical reasons. This failure is particularly important on our sentimentalist framework, as it might threaten the reasons-responsive status of our emotions. I characterise this kind of depression as a change in our *normative outlook*, the perspective we take on the world that sensitises us to some (but not all) of our practical reasons. Understanding how our normative outlooks change offers more fine-grained descriptions of how our agential capacities can be injured and restored as depressions emerge and remit. These changes in our agential capacities will have a significant effect on how far our autonomy can extend in various dimensions, but crucially, autonomy can still be built on impaired agential capacities, as long as those capacities persist in some state and we satisfy the social conditions for autonomy. Thinking in terms of normative outlooks allows us to extend recognition to impaired agential capacities, securing more autonomy than depressed people often enjoy.

I begin by outlining the sentimentalist position offered by Christine Tappolet (2016) which demonstrates our capacity to act autonomously when acting from emotion. I introduce futility thinking as an example of a drastic change in these capacities. In order to track the changes in these capacities, I develop my

account of our normative outlooks, as the particular range of sensitivities to reasons that we enjoy at any given moment, and which are susceptible to changes imposed by affective shifts, hunger, tiredness, and a wide array of other variables. This addition allows us to track not only improvements or impairments in our emotional reasons-responsiveness, but also changes in the sets of reasons to which we are sensitive. I then offer an account of regulative guidance that explains the specific kinds of error that depressive normative outlooks are prone to make, and which forecloses the objection that such normative outlooks are mistaken in a way that eliminates the reasons-responsive status of emotional motivation. I conclude by considering the effect futility thinking has on the shape of our autonomy, and the kinds of recognition required.

§1.2 - Sentimentalism

We are often brought to action by an emotion. The sentimentalist position holds that we (frequently) act *for reasons* when we act because of an emotion. The things we do because we are angry, or excited, or sad are not mere expressions of an arational state. Anscombe invites us to consider a particular application of the question '*why* was that thing done', an adequate answer to which is one that highlights the agent's motivating reasons, (2000, §5). Emotions are, on sentimentalist theories, rationalising answers to the Anscombian *why* question. To ask why someone shouted, in Anscombe's sense, might be answered by saying that they are angry at being ignored. Their reason for shouting is not, in this case, to attract attention, but rather their reason is that *they should not be ignored*. It is wrong and slighting for them to be ignored, which makes them (justifiably or otherwise) angry. Their anger at being ignored is their reason for shouting.

However, we might also respond to the why question by noting that angry people are prone to shout or slam doors. Indeed, we can present people's behaviour as the expression of emotions in a way that obscures their reasons for being angry, sad, or excited. Doing this can be a way of disempowering some agents by casting them as the vehicle by which an emotion, or a fiery temperament (presented as arational or irrational), is acting on the world. Explaining why Alice shouted with reference to her

emotions can answer Anscombe's question if we consider what she is angry about, or it can disempower Alice, if we suggest that her anger is an instance of 'women being emotional', deploying a stereotype attached to Alice's social identity (Mitova 2020, 709).⁷ If we are to empower people when they act emotionally, we must enquire after their reasons for shouting. The sentimentalist thesis invites us to do just this. Given BoRD is marked particularly by atypical emotional experiences, this kind of pathologizing non-recognition is a pervasive threat to bipolar autonomy.

Tappolet offers a persuasive theory in this vein. She holds that emotions function as quasi-perceptual experiences of practical reasons. Acting from emotion expresses our agency when so acting is acting in response to reasons. We respond to reasons insofar as our being motivated by emotions meets two requirements. The first is that our emotions are generally accurate in tracking reasons, and the second that we retain *regulative guidance* over our action (2016, 176). The first requirement is largely met by the analogy between emotional experiences and perceptual experiences. Taking emotions to be quasi-perceptual experiences of practical reasons is sufficient to mean that we can, to greater or lesser degrees of success, track our practical reasons by attending to our emotional responses.⁸ The second requirement makes it the case that our emotions do not count as motivating us without the kind of regulation that turns a tracking system into a responsive system, (173). She argues that all that is needed to make a tracking system into a responsive system is that a kind of regulative guidance prevents the tracking system from going too far astray.⁹ As regards emotions:

[T]o be committed to regulative guidance does not require more than cultivating and exercising the habits of reflective self-monitoring, what this entails concerning actions is simply that these habits of self-

⁷ These kinds of answers are not contradictory. In some cases they will be true at different resolutions. We empower people not by correctly identifying their reasons, but by choosing to speak about them in the register at which their reasons could be identified.

⁸ The perceptual analogy involved in the sentimentalist position remains controversial. Recently, Mitchell has emphasised the limits of the analogy, (2021, 95). That emotion and perception share certain features can be explained by their belonging to the same class of mental states, and does not give us reason to think of emotions as being perception-like in any additional ways. This gives us reason not to extend the analogy too far when trying to understand emotions on their own terms. However, for our purposes we need only understand emotions as perception-like to the extent that we can speak about their role in motivation and moral epistemology. For this reason, I continue to refer to emotions as quasi-perceptual experiences of the values of things, whilst acknowledging the force of Mitchell's argument as it pertains to the theorising of emotions as such. For previous contributions to this debate, see Cowan 2016, and Carter 2020.

⁹ She adopts this view of regulative guidance from Jones, (2003, 194).

monitoring are in place, so that they can be cultivated and exercised.

(Tappolet 2016, 176).

Regulative guidance is needed to turn a tracking system into a responsive system as the system's outputs need to non-accidentally correspond to the reasons that make them appropriate. Our emotional habits might be more or less effective at representing practical reasons to us, and in keeping with the Aristotelian tradition (in which sentimentalism is rooted), we always have some space to cultivate better habits. To let our emotions motivate us does not detach us from our reasons as long as we are keeping an eye on how well our emotions are tracking those reasons. In short, emotional motivation expresses agency as long as our emotions track reasons well enough that we can let them motivate us, while only stepping in to use other reasons-responsive systems when we have reason to believe something has gone astray.

Before continuing, let me briefly clarify how I understand reasons. I follow Alvarez (2010) who argues that reasons are facts viewed from a particular evaluative perspective. In light of a good, a certain fact is a reason-to- φ , (the good makes φ -ing right or appropriate). To φ might be right in more than one respect, as more than one good makes it appropriate. We can also view facts in the light of bads, negative evaluations which also reveal normative properties of the fact. Only facts can be reasons, and so falsehoods can only be 'apparent reasons'. As I am interested in the situations in which reasons become, or cease to be, discernible to us, I will sometimes talk about reasons becoming apparent to us. This does not mean that a non-reason is what is in question. Once we evaluate something as a good, in light of which we can see reasons, the most natural way to express this is to say that a reason has become apparent. So, I shall refer to apparent reasons, or reasons becoming apparent, without remarking on whether they do, in fact, justify φ . For clarity, I shall use the term *putative reasons* to pick out falsehoods which are mistaken for facts and so mistaken for reasons. Whether reasons are legitimate makes a difference for whether some action is justified, but it does not make a difference for the process by which we handle reasons.

§1.3 Depression and Futility

Depressions seem to involve a change in our moral perception. Depressions involve, and are arguably constituted by, dispositions to emote in ways that we would not otherwise emote (at least not with that consistency in tone). One kind of depression, which I term *futility thinking*, involves a shift towards boredom, indifference, and despondency. If emotions are perceptual experiences of practical reasons, how should we understand cases in which emotions cease to report that there *are* any reasons? Can our emotions report that there are no reasons to do anything in particular, (which thesis, Kahane terms *practical nihilism*, 2017, 330)? I contend that they can, and in a particular kind of depression, frequently do. I will offer examples of where this futility thinking seems to be at work, before exploring how I take it our emotional perception operates.

§1.3.1 Biographical Examples

We see futility thinking at work in the testimonies of depressed people across the literature. James' lectures on *The Sick Soul* include a wealth of instructive passages. He cites Tolstoy:

And yet I could give no reasonable meaning to any actions of my life. And I was surprised that I had not understood this from the very beginning. My state of mind was as if some wicked and stupid jest was being played upon me by some one. One can live only so long as one is intoxicated, drunk with life; but when one grows sober one cannot fail to see that it is all a stupid cheat. What is truest about it is that there is nothing even funny or silly in it; it is cruel and stupid, purely and simply. [... This is] the literal incontestable truth which every one may understand. What will be the outcome of what I do to-day? Of what I shall do to-morrow? What will be the outcome of all my life? Why should I live? Why should I do anything? Is there in life any purpose which the inevitable death which awaits me does not undo and destroy?

(This is James' own translation of passages from a French edition of 'A Confession', James, 2002, 123)

James is concerned with the ways in which our emotional life can (mis)represent the value of the world to an agent, even an agent "as powerful and full of faculty as [Tolstoy] was", as being "so ghastly a mockery." (125). James goes on to offer vignettes of similar depressions. He quotes the Rev. Henry Alline writing:

Everything I saw seemed to be a burden to me; the earth seemed accursed for my sake: all trees, plants, rocks, hills, and vales seemed to be dressed in mourning and groaning, under the weight of the curse, and everything around me seemed to be conspiring my ruin. My sins seemed to be laid open; so that I thought that every one I saw knew them, and sometimes I was almost ready to acknowledge many things, which I thought they knew: yea sometimes it seemed to me as if every one was pointing me out as the most guilty wretch upon earth. I had now so great a sense of the vanity and emptiness of all things here below, that I knew the whole world could not possibly make me happy, no, nor the whole system of creation. When I waked in the morning, the first thought would be, Oh, my wretched soul, what shall I do, where shall I go?

(Aline 1806 25, in James 2002, 127)

Solomon (2001) offers a similarly rich array of ways to be depressed. One of note will round out our exposition:

Lolly began to experience major depression. "I'd had a job but I had to quit because I just couldn't do it. I didn't want to get out of bed and I felt like there was no reason to do anything. ... I wouldn't get up to eat or anything. I just didn't care. Sometimes I would sit and just cry, cry, cry. ... I had nothing to say to my own children." (Solomon 2001, 341)

This is not the only time that Lolly speaks explicitly in terms of reasons to act (340), crystallising something about the existential state that Tolstoy and Aline commented on. The evil facts of the world, that it is fated to involve death, and that whatever meaning God can confer upon it he has kept in heaven, does not only take the wind out of motivational sails, but seemingly out of normative ones. In such a world, it is not clear to these agents what sufficient reason to do something or to live some way would look like. Lolly required reasons to do things such as work at a job or interact with the non-abusive members of her family, but could not find them.

Whilst I do not trade in terms of desire, something like the futility thinking I outline here has been noted by theorists who oppose desire-satisfaction theories of well-being, (Tully 2017, Spaid 2020).¹⁰ Tully presents 'complete conative collapse' as an actual, (rather than theoretical) counterexample to desire-based theories of wellbeing. From Frankl (1986, 173) he takes an account of the totalising apathy that many prisoners in the Auschwitz prison camp suffered. This apathy was a state wholly without desires which, in extremis, became a predictor for very imminent death. Tully takes this to be the extreme,

¹⁰ Cf Mariqueo-Russell (ms).

globalised, form of what happens when motivational anhedonia features as a symptom of depression. Motivational anhedonia involves impairment of anticipatory pleasure, of ‘wanting’ as opposed to ‘liking’, (Tredway & Zald 2011, 540). The activation we normally feel when an attractive option is present is missing, injuring the ‘incentive salience’ of our options, (Sherdell et al 2012). The interesting feature is that depressed people are often susceptible to the same amounts of pleasure when experiencing these goods. ‘Liking’ an experience ceases to be relevant to whether we perceive much reason to pursue that experience. Anhedonic depression can present not with the cessation of pleasure, but with the disappearance of apparent reasons to pursue those pleasures. Agents experiencing complete (or near complete) conative collapse, if they are as articulate and reflective as those cited above, might readily endorse the claim that they have no reasons to do anything in particular.

§1.3.2 Futility Thinking

Let us characterise futility thinking as follows: futility thinking is a disposition to represent the world to oneself as containing few or no ends that are particularly worthwhile and containing little that is of value. It involves dispositions to characterise the normative landscape as more or less barren. Futility Thinking is not the stalling or shutdown of an evaluative system, but rather is a grim report from it. The answer to all questions of value is some variation on ‘it’s pointless’. Futility thinking posits certain evaluations to be accurate as long as the values are close to zero or are negative. The agents above seem, to varying degrees, to still be invested, in questions of value, however dissatisfied they are with the answers they find. Indeed, this is an unpleasant situation to find oneself in precisely because we are scanning the normative landscape seeking good reasons, but recognising none.

Which dispositions make up this picture? When prevailed upon to take up some task or endeavour, futility thinkers are disposed to refuse, to not be persuaded that the end is worthwhile and to consider reasons given to be specious and a little insulting. Futility thinkers are disposed to seek paths of least resistance, to let ongoing projects slide, and are ill-disposed to take up new projects. We saw these tendencies in Lolly’s indifference and Tolstoy’s sense of pointlessness. Agents who come to hold a futility perspective often look back askance at their previous values (‘I cared so much about x , but that was stupid of me’) as exhibited by Tolstoy’s contempt for the “stupid cheat” he was subject to, and for himself, for falling for it. This grim outlook can extend over one’s own interests. An agent can recognise

that eating well and ensuring they have enough sleep are likely to secure their physical health without concluding that they have particular reason to do these things, as Lolly did. The agent casts themselves out of the moral community by treating their interests as considerations that fail to speak in favour of action.

Unlike cases of *accidie*, futility thinkers do not report some values that fail to move them, but rather repudiate values that they might previously have endorsed.¹¹ Consequently, the futility thinker's inertia stands in the same relation to their normative outlook as does the action of more enthusiastic agents. Futility thinking rationalises inertia the same way practical reasons rationalise action. If we flip the Anscombian question to 'why don't you do something you will enjoy' we get the quite intelligible answer in 'there's no point'.

§1.3.3 Futility Thinking and Agency

Depression is a shift in an agent's mood state, constituted by a change in the emotions the agent is disposed to experience. This change leads them to build a picture of the normative landscape that looks barren, and to endorse something that looks like practical nihilism. We must now ask what has happened to the agent's moral perception when they slip into this mindset. Our answer to this question will determine what we have to say about depressed agents' autonomy. If this change in our emotional dispositions means our emotions no longer compose a reasons-responsive system, then depression has wholly arrested the agency which is the foundation of an agent's autonomy. What is happening to our reasons-responsiveness when our emotional dispositions change in this way, and when they recover, is therefore foundational to our broader investigation.

What we seem to have is an emotional state that reports the absence of reasons. Agents whose emotions had previously been well tuned to track practical reasons find that system reporting to them not only that there are no good reasons to φ , but that any putative reason is meagre and rather

¹¹ I characterise the broad category of depressions in which agents continue to report values or cares that they feel should motivate them, but which fail to, as *accidic*. See Watson (1975, 744), Tenenbaum (2003), Calhoun (2009), Gorman (2021).

insulting. We could understand this simply by noting that depression has struck, and so the agent's reasons-tracking system has been interfered with. Their futility thinking is an expression of a social, spiritual, or neurochemical phenomenon that adequately explains it. This is to opt for a disempowering explanatory approach to the *why* question. We do not take Lolly's word as moral testimony when she tells us that there seemed to be no reasons out there which would make getting up and going to work appropriate. Accepting this interpretation consigns depressed people to impaired autonomy by refusing to recognise that there is an independent, and somewhat disquieting, perspective on the normative landscape at work here. If we are to seek the agential perspective, according to which Lolly and other futility thinkers might autonomously endorse this nihilistic perspective, we must determine how her moral perception is working, or not working. Whilst there are certainly injuries to agential capacities, we must distinguish the injury to autonomy that stems from impairments in these capacities, and impairments to autonomy that stem from non-recognition of these capacities. Whilst futility thinking is a species of unsuccessful agency in the sense that it involves a mistaken view of the normative landscape, it is still the exercise of capacities to interrogate facts for their normative properties, and to compose a picture of the world that highlights its normative properties. There is an exercise of some agential capacities at work here, that is due recognition for what it is.

§1.4 Normative Outlooks

A *normative outlook*, as I use the term, is the means by which an agent resolves the various facts (and putative facts) they encounter into a landscape of reasons with which they might interact. Agency is not possible without an outlook that focuses one's attention on a manageable set of reasons. To extend the perception analogy, if goods and bads are wavelengths of light which reveal the normative properties of things, then a normative outlook is the range of wavelengths an agent has the eyes to see. We all experience changes in our normative outlooks when we are too tired or hungry or grumpy to appreciate certain reasons. For some people, the more significant of these changes might be infrequent and short-lived variations in what is otherwise a succession of extremely similar normative outlooks. People who live with BoRD, however, will repeatedly and predictably experience very drastic changes in their normative outlooks, where they are insensitive to a range of reasons or extremely sensitive to reasons

they would usually give little heed to, for weeks or months at a time. The majority of neurotypical agents lie between these extremes, being generally sensitive to a stable family of reasons, but for the usual variations due to weariness, drunkenness, hunger, and the like.

A normative outlook is composed of the range of normative properties an agent has the eyes to see and admits of high degrees of change across even quite short periods of time. It is distinct, therefore, from the sets of concerns that might constitute a *practical identity* or *deep self*, as is popular in some accounts, (Frankfurt 1998 11, Van Willigenburg & Delaere 2005, Shoemaker 2015, 44-56, Killmister 2017 19-33, Gorman 2021). Normative outlooks could not do this work as it is their susceptibility to change that is what makes them theoretically useful when considering affective shifts such as the onset of depression. Rather than becoming different people because we cease to care about what we cared about or revise some judgement that gave shape to our identity, we can be much the same person now, occupying a new normative outlook. There might be some change in the description that we take ourselves under as a result of such shifts, but we are not new people each time.

A person might have emotional dispositions that are generally well tuned to tracking reasons, supported by regulative guidance, but which locally fail to track adequately. This can simply be the amount of 'give' that is implicit in a disposition, (dispositions hold even when they are not expressed). Sometimes, however, these failures amount to a change in their normative outlook. We have observed that changes in our moods involve changes in our emotional dispositions, thereby (per sentimentalism) changing our sensitivity to reasons. The following account explains how these changes come about at the subpersonal level.

§1.4.1 Salience and Valence

Two elements that are crucial to composing a normative outlook are salience and valence. The salience-granting work of emotions supports deliberation by marking objects for normative attention (Damasio 1996 173). Any fact might disclose a reason if viewed in an appropriate normative light, but as we do not have space to deploy all those reasons in our deliberation, we must be selective. Failure in this kind of selectivity, which has significant consequences for agents' efficacy, is associated with lesions in the

parts of the brain understood to be crucial to emotional processing. Participants who endured ventromedial frontal lobe injuries were generally ineffective in playing a lifelike gambling game, Damasio argues, due to impairment in their ability to mark as salient particular facts that were normatively significant for their practical purposes, (212-222). Our emotional systems function to mark for salience those facts that are normatively significant for us. Reasons are ubiquitous and so we need some way of narrowing our attention to the ones we most wish to, and have most need to, consider.

Once an emotional reaction attaches to some fact, it is marked for salience, we require a normative light to reveal its normative properties, which we can respond to in deliberation. This is where valence becomes relevant. The normative properties, or reason-status, of a fact is only apparent when illuminated by a good or a bad (following Alvarez, above). We can be motivated by seeing some fact in the light of a good or of a bad. These are reasons to do, to feel, or to believe, or are reasons to avoid, to forbear, or to decry. Certain emotional experiences such as despair are perceptions of reasons illuminated by bads. The difference here is between the positive and negative valence which these perceptions involve. Thus, emotional experiences must admit of valence which can track the evaluation that illuminates more specific normative properties. As Mitchell has it:

“emotional experiences all involve positively or negatively charged intentional feelings as felt favourings or disfavourings, as a primitive non-doxastic ‘yes’ or ‘no’ which ‘comments’ on the evaluative standing of the particular object.”

(Mitchell 2021, 107)

As Alvarez presents it, there is more than one good that we might deploy to illuminate the normative properties of some fact, (2010, 14-5). We can reflect on whether we love a novel because it is *kind* or because it is *beautiful*, each to whichever degree. Thus, salience-marking and simple valence are not sufficient to exhaustively track our practical reasons (failure to distinguish moral from aesthetic reasons being only one example). Certainly, more is involved in emotion than only these two features, as curiosity, wonder, and gratitude all involve similar patterns of salience and valence. I will not explore this point further as we do not require a complete theory of emotion for present purposes. All we require is an understanding of the features that make emotions sufficient to track practical reasons, for the purposes of becoming a reasons-responsive system, and which are involved in the changes in experience and motivation that depressions bring on. Futility thinking can be expressed by a wide range

of discrete emotional states, but are united by their neutral (when numb) or negative (when despairing) valences.

The range of emotional experiences we are disposed to have in any given situation establish for us the basic grammar we will use to describe the normative landscape. The range, number, and diversity of objects worth interrogating for their normative properties is set by what is marked for salience. The normative valence of those properties is set by whether we are disposed to emote positively or negatively towards some object. It is in this way that our emotional dispositions, at any given moment, sensitise us to a range of normative considerations. Crucially, they do not sensitise us to *all* considerations. Our emotional dispositions can shift quite drastically on good or bad days, due to different moods, or whether it is before or after lunch.

There is an important nuance to capture about valence. Charland has argued that the valence of some affect is indeterminate until the agent experiencing it turns inwards and examines their affect. There is some quale, something it is like to be feeling guilty about something, but until we recognise that we feel guilty, and fix its hedonic tone by attending to it, it is without valence. Affective experience is altered simply by being attended to, and so:

Valence... is a dynamic relational evaluative phenomenon that emerges out of the interaction of attention with first-order phenomenology. In other words, affect valence is neither purely intrinsic and “found,” nor purely extrinsic and “constructed.” It is enacted.

(Charland, 2005, 247)

Our emotions will mark particular facts for our normative attention, however before we can deploy them in practical deliberation, we need to establish their valence. This isn't primitive in the experience. Why, therefore, would Tolstoy or Lolly or any other futility thinker valence their affect in this way?

Ratcliffe has argued, persuasively, that various sorts of depression are characterised by a change in deep, *existential feelings*, which he takes to be pre-intentional states which themselves lack object or intelligible content, but which scaffold the set of intentional states available to us (2010, 2013, 2020).

Various kinds of depression, despondency, or grief are, on Ratcliffe's account, underpinned by

existential feelings of deep guilt (2010), or an inability to recognise certain futures as being possibilities *for oneself* in a properly engaged way (615). Valencing an experience in the way Charland outlines requires applying a particular kind of intentionality to that experience, and so will be confined by the way the existential feelings at work scaffold, or do not scaffold, one's intentional states. Some change in existential feelings will then make it difficult or impossible for us to valence positively those objects (facts, memories, situations) that are marked for salience. If Ratcliffe is correct, then affective shifts such as the onset of depressions involve a change in the existential feelings that provide us with the range of valences we might attach to the feelings that are marking objects of our normative attention for salience. Depressions, particularly when they amount to futility thinking, I contend, are the result of shifts in our existential feelings such that we cannot positively valence the affects that make up our emotional perception of practical reasons.

Futility thinking involves a loss of the interpretive tools we would use to confer positive valence on our affect. Without this positive valence, we only see reasons in the light of bads. Depressions which are not marked by sadness as much as they are marked by numbness deprive us of negative valence as well as positive valence, and so rendering facts into practical reasons is vastly more difficult for an agent. Under normal circumstances, our emotional experiences mark some facts as salient, and we ascribe valence to these emotions which provides the normative light. Our emotions, therefore, disclose facts to us under a normative description, making the reasons they constitute accessible. Changes in our existential feelings will change what emotional experiences we have and so what normative outlooks we can compose at different times. Depressions which are marked by pervasive loss of positive affect, or of any affect but numb neutrality, deprive us of access to whole categories of reasons at a time. Futility thinking is what occurs when the emotional experiences we rely on can no longer reveal reasons to us in the light of a good, or eventually, cannot reveal reasons to us at all. The agent's normative outlook can report only practical nihilism: that there are no reasons to do anything in particular.¹²

¹² We must be cautious not to overgeneralise. Even when depressed, we can often still valence *some* of our affects positively. The fact that a person laughs at some jokes does not mean that their depression is less severe than it was at some other time, when they did not laugh at anything. Deliberation is still possible using any reasons in the light of goods that remain apparent to us, or using reasons in the light of bads. Futility thinking is probably expressed in a general tendency in agents' normative outlooks, to reject reasons as inadequate or as merely putative. It is unlikely, in most cases, to be a totalising phenomenon.

§1.4.2 Reasons-Responsiveness in Futility Thinking

The account so far holds that we act for reasons insofar as we are motivated by emotional states that adequately track reasons, and that the cases where our emotions report to us that there are no sufficiently good reasons, or that the only reasons available are reasons to rue or decry, can be explained with reference to changes in how our emotional states mark salience and confer valence. The account explains depressive inertia by suggesting that agents engaged in futility thinking compose a normative outlook with which they simply cannot recognise many facts as being sufficiently good reasons to φ . This range of practical reasons accessible to the agent is greatly diminished and so there is little to do. This also explains the grim set of descriptions the agent chooses to give to matters in their life, in cases where reasons in the light of a bad are still apparent. Alline's sins were brought to mind every time he met someone. Facts he was already disposed to valence negatively were all that his emotional life would mark for salience. His antipathy towards his self is predictable if the only things about himself that seem significant are the ones he is already disposed to valence negatively. In cases where few to no reasons at all are apparent, this predicts precisely the inertia we see so often in depressed people, as well as rhetoric of futility and pointlessness. Pervasive numbness deprives us of anything which might hold our attention under a normative description, and so no consideration seems to merit more than an indifferent shrug. Importantly, this is a warped picture of the normative landscape.¹³ There are a great many more reasons to act than the seriously depressed agent is sensitive to, so their depression gives rise to a failure in their reasons-tracking capacities.

The crucial element in this account is the normative outlook, understood as the kind of sensitivity to reasons that an agent has at any moment. Normative outlooks are subject to change on the bases of even quite small changes in our emotional dispositions (such as having missed lunch) or quite large ones, such as the affective shifts involved in bipolar disorder. If the normative outlooks we can compose are heavily skewed as they are during bouts of futility thinking, then our emotions will fail to track reasons to high degrees of success. For our emotions to feature as part of our agency in these contexts, we must fall back on regulative guidance. We must be committed to "habits of reflective self-

¹³ As noted above, I take a factive view of reasons, and so reject any reasons internalist positions which would hold that futility thinkers have fewer practical reasons than they used to.

monitoring" (Tappolet 2016, 176). The excerpts above seem to indicate that, to varying degrees, these commitments are present. Tolstoy is the most articulate on this point, evidencing a deep investment in the appropriateness of his nihilistic normative outlook. This was the tendency to look back askance at one's previous values.

This tendency expresses what Westlund has called self-answerability: the disposition to consider how one would justify one's view of the normative landscape to appropriate interlocutors (even if those interlocutors only exist in one's moral imagination, 2009, 36). This disposition is crucial to autonomy as it prevents "passivity in the face of one's commitments" (34), but also would seem to satisfy the commitment to regulative guidance. To be self-answerable, as Tolstoy certainly was and as Alline and Lolly seem to have been, is to have a habit of considering whether interlocutors might be able to offer objections sufficient to shift one's view of the normative landscape. Our futility thinkers seem to have one eye out for the 'red light' that Tappolet holds is supposed to go on when our reasons-tracking system fails, (2016, 117). There are three possibilities as to what is happening here. The first is that the red light is going on, but is being ignored, meaning that they are not really committed to regulative guidance. The second is that the agent is looking for the red light but it is not going on. The third is that the red light is going on, but is being misidentified as something other than a warning that one's emotions are no longer reliably tracking reasons. The second and third options both mean that the agent remains committed to regulative guidance, but that this commitment is proving ineffective, and seem to be better candidates to capture the tendency to look back askance that we observed.

What does this mean for the depressed person's agency? Certainly there is damage to their agential capacities in that their emotions can no longer reliably track practical reasons. They remain committed to regulative guidance, but this commitment is not ultimately effective, for reasons we shall explore in the next section. The question of whether their emotions still constitute a reasons-responsive system therefore turns on whether 'regulate' is a success verb in this context. Whilst it is tempting to hold that it is, which would suggest that the depressed person's agency is arrested during the period when their depression interferes with their reasons-tracking and regulation, I argue that it is not. If regulation were a success verb, then the only way to be an agent would be to appropriately respond to our practical reasons. Yet it implausible to suggest that our agency is not at work when we do the wrong thing because we misidentify, misweigh, or are akratically distracted from, our winning practical reasons. Our

commitment to regulative guidance makes a tracking-system non-accidentally get questions of our reasons right most of the time, but the standard cannot be that it always gets our reasons-tracking right or there would not be space for agents to act in violation of their reasons. Such violation would be agency-defeating on the grounds that the commitment to regulative guidance did not successfully correct them on this occasion.¹⁴ Thus, a depressed person's agency persists, albeit impaired, when in a state of futility thinking. But precisely why regulative guidance has been ineffective in this context remains important for the agent's autonomy. The next section will offer a theory of regulative guidance in these kinds of depression.

§1.5 Regulative Guidance

One way regulative guidance might be exercised is by recognising that one is depressed, and so become sceptical about one's normative outlook insofar as it is composed by emotional reactions. Plausibly, this is what we do when we practice cognitive behavioural therapy and its variations. Fortunately, more than one reasons-tracking system contributes to the normative outlook we occupy at any moment, and so we can throw to another reasons-tracking system when needed. Our moral judgement is the main alternative and is one we might use to correct for errors in our normative outlook occasioned by depressive futility thinking. More importantly, I suggest, this is the system we are most likely to rely on to signal that something is amiss with our emotional reasons-tracking in the first place. It is disconsensus between our reasons-responsive systems that flashes the red light that we are looking out for when we are committed to regulative guidance. Having established the way in which shifts in our emotional dispositions change the normative outlook with which we are able to track some portion of our reasons, this section will extend the account to explain why it is that the depressed agent's commitment to regulative guidance is unsuccessful in the ways we saw in the previous section.

¹⁴ For a fuller discussion on this point, see McGeer & Pettit, (2015).

§1.5.1 Moral Judgement

When a person ruminates on matters of value, and judges some end or other to be worthwhile, they form beliefs about value and about which reasons command them. Watson has defended judgement as “an original spring of action”, insisting that we can judge our way to (at least moral) motivation. He tells us:

It is because valuing is essentially related to thinking or judging good that it is appropriate to speak of the wants that are (or perhaps arise from) evaluations as belonging to, or originating in, the rational (that is, judging) part of the soul; values provide reasons for action.”

(Watson 1975, 17).

I differ in not thinking that judging is essential to valuing, but merely sufficient. Consider the devout Kantian who is persuaded by a Kantian line of reasoning that the moral patienthood of all agents is equal, and so comes to value others no less than themselves. Perhaps some passion for their own wellbeing was the starting point from which they came to value anything, but their appreciation of Jemima’s value as a person was reached by reasoning, and their actions towards Jemima are motivated, in an upstream way, by that evaluative judgement.

Judgements about value admit of valence. We judge some object to be good or bad, and some fact to justify, require, or forbid some action. That we have formed a belief on something indicates that it was marked for salience at some point – at least enough salience for us to judge about it - although the subject matter may cease to be salient to us without us ceasing to hold the particular belief we had formed. Judgements of this sort create beliefs that cast some fact as a reason to, or to not, φ . In this way, judgements sensitise us to particular reasons for the future, more directly than emotions do perhaps, by committing us to take these or those considerations as reasons on pain of inconsistency. Our overall sensitivity to reasons is established not only by our emotional dispositions, but also by our judgement. Insofar as judgement issues in beliefs, we are disposed to be sensitive to the normative properties our evaluative judgement commits us to track.

§1.5.2 Regulative Guidance as Sustained Pressure

It seems that our reasons-responsive systems are able to check, temper, and discipline each other. Emotions can drive our actions, but we can pause and consider whether we trust those feelings to direct these actions, in this moment. This is the core of cognitive behavioural therapy techniques and is valuable if we do not trust certain emotions. Similarly, emotion can check our judgement. Reasoning can take us to conclusions that feel wrong, despite our acceptance of every premise and link, as they pick up on something that our consideration has missed. Huck Finn judged, based on what he took to be reliable moral testimony, that slaves were property and that not turning slaves in was as bad as stealing. Nonetheless, his fondness for Jim prevented him from acting on those beliefs.¹⁵ The pressure applied by disconsensus between his reasons-tracking systems is what signalled to him the need to exercise regulative guidance over the system.

But the point of disconsensus is not the only instance in which these systems interact. Indeed, they are constantly interacting. Our emotions push and nudge our judgement around by holding our attention on the facts we had taken to be irrelevant, marking them for continued salience so much that we repeat the judgement process. Judgements about what is valuable can, the more we are reminded of the conclusion we came to, dispose us to emote in relevant ways. These systems offer the red lights for each other that allows us to engage in regulative guidance, but this is an expression of their general tendency to pull each other, for better or worse, into coherence.

Certain kinds of moral learning occur when our emotional reactions apply continual pressure on our judgement. Jagger (1986) highlights the role ‘outlaw emotions’ play in the epistemology of justice. ‘Outlaw emotions’ are more commonly experienced by subordinated people, and it is these emotions that represent the injustice of one’s subordinated position. “Anger” we are told “becomes feminist anger when it involves the perception that the persistent importuning endured by one woman is a single instance of a widespread pattern of sexual harassment,” (166-7). She writes:

Outlaw emotions stand in a dialectical relation to critical social theory: at least some are necessary for developing a critical perspective on the world, but they also presuppose at least the beginnings of such a

¹⁵ See Arpaly (2002, 76) for a fuller discussion.

perspective.

(1986 166-7)

The process of coming to consciousness (in this case, feminist consciousness) of the injustices which masquerade as 'simply the way of the world' can be achieved by reasoning, but often only after the moral emotions have kept up sustained pressure. Jagger's emerging feminists are exercising regulative guidance by recognising the discomfort this sustained pressure creates as reason to doubt that their judgements about how they should be treated, deeply influenced by patriarchal norms as those judgements are, are reliable.

Schematically, regulative guidance is what occurs when Huck Finn continues to believe that slaves are property and so neglecting to turn an escaped slave in is theft, yet resolves to 'remain a bad boy'. His affection for Jim is sensitive to Jim's independent value as a human being, and Huck decides to trust in that instead. Rather than being a case of slow attrition, Huck is faced with a choice between two reasons-tracking systems, and defers to his emotions. But I take it that the slow moral learning process is a more commonplace and lifelike example of regulative guidance at work than the schematic case. Regulative guidance is expressed by habits of reflective self-monitoring, which can operate deeper in the background and be more distributed across time. Openness to change one's mind, and the recognition that one's emotions offer legitimate moral testimony against which one's judgements can be checked, expresses the commitment to regulative guidance that makes moral judgement into a responsive system. We can explain the failure of regulative guidance to be successful in cases of futility thinking with reference to this sustained pressure that these systems apply to each other.

§1.5.3 Misguided Regulative Guidance

We have seen that continued pressure applied to one's normative outlook by emotional perception of reasons can be a source of moral knowledge. The emerging feminist learns that her treatment is unjust, nor merely uncomfortable, by being open to the influence of her emotions. However, this process can cut the other way, and be the reason that we lose moral knowledge. It has been controversial whether it is indeed possible to lose moral knowledge, but McGrath has recently and persuasively argued that moral knowledge can be lost when it is defeated by rebuttal, (2019). Our knowledge is defeated when rebuttal undermines our justification for our moral belief to the degree that it is no longer knowledge.

Such rebuttals are found, she argues, when two people disagree on a point of morality. Two premises hold for McGrath's argument to work. The first is that the testimony of other people can provide us reasons to believe a moral proposition p . The second is a conciliatory view of disagreement, according to which the fact another epistemic agent disagrees with us with respect to p gives us reason to lower our credence in p . McGrath requires only the most minimal conciliatory view, and one that follows quite neatly from the view that moral testimony is evidence. If some (qualified) person's testimony is reason to believe p , then it can also be reason to believe not- p , (176). If testimony is sufficient to be evidence for a moral proposition, it is as easily evidence for its negation. Rebutting evidence defeats our moral knowledge, the argument goes, when it accumulates to the degree that we are no longer justified in believing the moral proposition that is true and which we had knowledge of prior to defeat.

We do not need to weigh in on the debate concerning conciliatory disagreement, nor the debate concerning whether testimony is evidentiary for moral propositions. We need only note, as McGrath does, that the evidentiary status of some kind of experience is unchanged whether it is evidence for a proposition or for its negation. We have adopted the sentimentalist framework, according to which our emotional experiences are (quasi-)perceptions of the normative properties of things. These perceptions are reasons to form moral beliefs, of the sort that partly compose our normative outlooks. However, contrary emotional experiences are also evidentiary, and can rebut our reasons to believe that the world is rich with value and meaning. Continued pressure applied to our beliefs about value, coming from our perceptual experience of value, can undercut our justification and so costs us moral knowledge as McGrath describes. This pressure moves us to revise our beliefs about value in just the same way it did for Jagger's emerging feminists.¹⁶ The difference here is that the change in our moral beliefs costs us, rather than wins us, moral knowledge. Having lost these beliefs about value, we are no longer sensitive in the way we were to the reasons that these values would reveal. Our normative outlook, partly composed by judgements about value we had made in the past, is altered by the same process that we rely on to correct it under more favourable circumstances. It is this openness to one's own emotional experiences, trust that they do in fact report on the normative properties of the world, that expresses a

¹⁶ These are two distinct processes of course. One issues in a change in what we are justified in believing, and the other a change in what we actually believe. My argument requires only that the latter happen, though I suspect the former does as well (meaning futility thinkers can be mistaken without being irrational in their nihilism). Even if McGrath is wrong about what such evidence does to our justifications, I contend we actually respond in the way she suggests would be rational.

commitment to regulative guidance in the emerging feminist but is the reason that regulative guidance does not successfully correct the agent's reasons-tracking errors in futility thinking.

What does this mean for agency during such depressions? The agent remains committed to regulative guidance; in the habit of letting their emotions and judgements apply pressure to each other that will reveal and ameliorate disconsensus, but which can lead agent's astray if circumstances are not favourable. In futility thinking, the reasons-responsive system fails in a way reminiscent of moral miseducation. A legitimate kind of evidence (emotional perception in the case of depression, testimony in the case of miseducation) is, in its particular context, unreliable evidence. This is a failure of a reasons-responsive system to respond properly to the reasons that are in front of it, but it is not a failure of that system to be a reasons-responsive system. Returning to Huck Finn's contemporaries, it seems we must attribute their mistaken moral outlook to the operation of their agential capacities, despite their moral miseducation. To whatever extent we might limit the blame we ascribe to them, if we are to consider them agents at all they are agents who have composed a quite wrong picture of the way normative properties are distributed throughout the world. That they are so miseducated that they cannot recognise a family of reasons dooms them to be less virtuous. It does not relieve them of the status of agent. It seems that we must be autonomous agents even when regulative guidance fails to correct us. Futility thinking impairs our agential capacities, costs us moral knowledge, and leaves us less able to find and respond to value as it exists in the world, but this impairment of our capacities does not render them non-agential. They are malfunctioning agential capacities, rather than being no agential capacities at all. We shall consider the consequences this has for depressed agents' personal autonomy in §1.7. For now, I turn to an adjacent case of depression, in order to demonstrate independent support for this account.

§1.6 - Accidic Depression

Depression is well-trodden ground in moral psychology and so futility thinking is not the only sort of atypical agency worth accounting for. Stocker highlights a particular pain of depression, noting that "a

frequent added defect of being in such “depressions” is that one sees all the good to be won or saved and one lacks the will, interest, desire or strength”. (1979, 774), and Gorman suggests that some “volitional gap” lies between a depressed agent’s deepest cares, which they continue to avow, and their motivation, (2020, 227). Tenenbaum’s scholastic view (2003, 2006, a close relative of sentimentalism) raises this problem under the title *accidie*. On Tenenbaum’s view, desires (construed widely) are perceptual appearances of the good, and accidie occurs when an agent desires something but feels no motivation to pursue it. This is certainly a feature of many sorts of depression. The key distinction here is that the depressive inertia associated with futility thinking is not accidic. Futility thinking is the state Tenenbaum (2003) distinguishes in endnote 19: “In many cases, patients simply ‘don’t care’ about anything in such a way that it would be hard to say that they still value anything”. Depression can, but does not necessarily, eliminate or perceptions of the good. The account we have developed so far will gain further plausibility if it can make sense of accidic depressions, as well as the futility thinking we have examined so far.

Tenenbaum draws on Kant’s suggestion that the happiness of the vicious person fails to be good, (Tenenbaum 2003, 160). The goodness of happiness is conditioned by the happy individual’s character. Tenenbaum imagines an agent who secures a holiday to New York by betraying a friend, the knowledge of which fact prevents her from enjoying her holiday. This agent does not represent a holiday to New York as bad, but during her stay she represents herself as disloyal. The good of this holiday is conditioned by her character, or by how it was won. Consequently, whilst the holiday appears good to the agent, it does not appear good in a way which can be realised for her. Tenenbaum suggests that this explains typical cases of depressive accidie, where apparent goods are conditioned by facts such as the depressed person’s character. He takes purported values which fail to motivate as having this conditioned structure: ‘it would be good for someone to go to their friend’s birthday, but I’m such a rotten person’, or ‘I should write my novel, if only the universe weren’t so cold and indifferent’. This view assigns explanatory function to familiar feelings of shame and guilt associated with depression. One could perhaps render the vignettes in §1.3.1 as involving this sort of conditioning. Tolstoy highlights the fact that he will eventually die as a defeater for any possible reasons he has to act. He might actually value some things, but find that value conditioned by something he doesn’t have (immortality).

Tenenbaum emphasises that he takes people who are accidically depressed in this way to be committed to a conditioning relation, (2003, 168). The agent who states that she desires something but that she feels no motivation to pursue it can be made intelligible insofar as we ascribe such a commitment to her. How, then, should our model account for cases of apparent accidie? It is far from clear how this kind of complex conceptual content could be delivered by a perceptual, or quasi-perceptual experience of the value of things. Indeed, on Tappolet's account, emotional perception is non-conceptual, (2016, 16).

I suggest the following as an alternative explanation for accidic depression. Accidie appears as the interim stage in the process outlined above. Our sensitivity to various sets of reasons can consist in our emotional dispositions or beliefs we have reasoned our ways towards. However, as our emotional systems continue to report that nothing is valuable, pressure is applied against this judgement. Accidie occurs during the midst of this process, while beliefs about value are retained, but the perception of value is lost. It is what happens when we deliberate about what to do but cannot lift ourselves from bed to act, because our evaluative beliefs have outlasted the perceptions that helped to justify them. In a pedantic way, I disagree with Stocker that accidic depressions are ones in which "one sees all the good to be won or saved but one lacks the will..." (1979, 744). Rather, we do not see the good, we believe it is there, and struggle because what we are seeing is its absence. This process needn't always conclude with all out futility thinking. Sometimes we will recover from our accidie because our judgement can help re-sensitise our emotions to the value of things, (again, this seems to be what a great many talking therapies achieve).

The advantage of this understanding is that it retains specific attention to the emotional states that are distinctive (indeed, constitutive) of depressions, whilst omitting complex conceptual content such as conditioning relations. Indeed, Tenenbaum does not require that accidic desires are themselves appearances of goods as conditioned (156), but only that we are rationally committed to the goods being conditioned by the structure of our motivation. The objective here is to make accidically depressed agents intelligible as rational, which is virtuously charitable in some ways. It is a mistake, however, to expect that rationality is the standard of intelligibility we should always aim for. We are flawed, inconstant, and imperfect agents, and so if we are to recognise agency in each other, we must be able to

recognise sub-optimally rational agency at work. The account I have offered allows us to make not only individual reports of evaluation intelligible in context of a depressed agent's action (or inertia), but it also makes intelligible the dynamic process whereby depressions emerge and recede. We can recognise intelligible agency not only in reports in context with actions, but as parts of an agent's project of building and updating a picture of the normative landscape they inhabit.

§1.7 What of Autonomy

So far we have developed an account of our moment-to-moment moral psychology, which explains how emotion feeds into our deliberation under normal circumstances, as well as why we don't undertake to deliberate at all during periods of depressive inertia. This is an account of how agency works during depressions. We now must ask, what does this kind of agency do to the autonomy of someone who experiences depression, or lives with BoRD? As set out in the introduction, our agential capacities are the base upon which our autonomy is built, and so we must know how impairments of those capacities transmit up through our autonomy. We are also interested in what kind of recognition this kind of agency requires if the conditions for socially constituted autonomy are to be met.

§1.7.1 Agential Capacities

The futility thinker systemically mistakes reasons for non-reasons, and is prone to lie in bed, to give up on their projects, and to refrain from taking up new projects. More generally, very depressed agents tend to not see their reasons straight, tend to do very little, and tend to give up on very many of their existing projects. In this regard, they do not seem to be doing well as an agent. Whatever one accepts as the specific dimensions of autonomy, these tendencies will limit the extension of an agent's autonomy in at least some dimensions.

Futility thinking prevents one from responding appropriately to reasons, but lets one respond to putative non-reasons in a way that would be appropriate if one were not mistaken. This matters

because it tells us where the chunk is taken out of the futility thinker's autonomy. When our emotions fail to be reasons-responsive, it is usually because emotions present a non-reason as a reason, or misrepresent the relative weight of reasons. In futility thinking, it is because our emotions misrepresent reasons as non-reasons, and bring our judgements into line with this error. Thus, we become less successful in the core business of agents, in responding to reasons.

Yet this is not a comprehensive failure. The futility thinker can still engage in the higher level, though less core, task of agents, which is to resolve the world they encounter into a normative landscape. The futility thinker will draw a skewed picture of this landscape. However, the business of agents includes composing such pictures, and we can answer a wider scope version of the Anscombian question with reference to their (skewed) picture of the normative world. Agents act-for-reasons, and during futility thinking, forbear from action for want of reasons. As we saw in §1.5.3, the mistake the futility thinker makes is similar in important ways to the mistakes made by people who are systemically morally miseducated. Depression is more likely to reduce the extent to which we think someone is blameworthy for their error than is being morally miseducated, but the similarity is sufficient for us to continue to recognise this malfunctioning system as reasons-responsive to the degree that constitutes agency.¹⁷ The errors are agency going wrong, not something ceasing to be agency.

Some agential capacities are left at least somewhat intact. Whilst there are fewer reasons that the futility thinker has within their ken, the reasons they have (to mope, rue, or decry) remain reasons that they respond to. We are biased in our thinking about agency towards responses to reasons that issue in publicly visible actions with apparent practical consequences. However, another key agential capacity that we all have, and which continues to function when we are depressed, is the capacity to *appreciate* the value, or disvalue, of things. Depression injures our ability to track where value is and fixes our attention on disvalue (sometimes illusory disvalue), but when we are not mistaken as to the disvalue of something, we can continue to appreciate that disvalue in the relevant way. This agential capacity is

¹⁷ We tend to expect people to be critical about the value of their cultures, and so take this kind of moral miseducation as at most a partial excuse. If depression is grounded in shifts in the existential feelings which structure the intentional states we have available to us, as Ratcliffe suggests (2010, 2013), then there is limited space for depressed agents to engage in this kind of critical reflection. The matter is much more complicated than this, of course. We address critical reflection in §5.3.2, and responsibility in §6.

damaged, but not beyond functioning, in depression, as whilst we struggle to appreciate the value of things we cared about, we can and often do properly appreciate the downsides of life.¹⁸

As autonomy is an agency concept, our agency provides the base upon which our autonomy is built. The impairments to depressed agents' capacities which we have explored have consequences for that agent's autonomy. The inertia that futility thinking gives rise to limits the exercise of our agential capacities. On different specific theories of autonomy, this will limit how far autonomy extends in different dimensions. On MacKenzie's theory (2014), this tendency need not limit their self-governance (allowing that Tolstoy's narrative expressed a genuine practical identity), but it will significantly limit the degree to which autonomy extends in the self-authorisation and self-determination dimensions. Compare with Killmister's theory (2017): extension in the self-realisation dimension and probably the self-constitution dimension will be capped by this tendency to misread the normative landscape. The dimensions more involved with coherence than accurate reasons-responsiveness (self-definition and self-unification) need not be limited by futility thinking.

Whilst the project we are engaged in is neutral with respect to what we should take to be the specific dimensions of autonomy and how many there are, let us pause to consider both self-determination and self-definition, some variant of each being likely to feature prominently (either as a particular dimension of autonomy or as an otherwise important feature) on most theories. There being little or no reason to act in any particular way or to continue with one's commitments raises two threats to self-determination. The first is that one is less likely to carry out one's commitments effectively if one gives up on one's projects because the normative world seems pointless. The second, perhaps deeper threat, is that one will cease to make the determinations that are exercises of autonomy in this dimension. This is inertia rather than passivity. Autonomy is not arrested in this dimension in the way it is if we passively accept the evaluations of others, but is inertia that prevents autonomy extending any significant distance in this dimension. This is a qualified injury to autonomy, therefore, as it is also the self-determination dimension of autonomy that is injured by the 'lack of responsiveness to considerations

¹⁸ Our excessive focus on unpleasant facts might suggest that something is amiss without appreciating, but the capacity is functioning at least moment to moment. Someone who is excessively prone to anger might be correctly appreciating various injustices at any given moment, even if their patterns of attention make them less than virtuous.

that purport to challenge our current sense of the justificatory landscape' which characterises an autonomy-defeating passivity on Westlund's (2009, 34) account. Yet Tolstoy, and the schematic futility thinker at least, exhibit precisely the non-passive stance towards their own values that Westlund highlights. Tolstoy's resentment at the cruelty of the joke that seemed to have been played on him exhibits a deep investment in his sense of the normative landscape being correct. The futility thinker looks back askance, or looks at others with derision, because they are not indifferent towards their own normative outlook.

In contrast to self-determination, it seems that the self-definition dimension is left unlimited by futility thinking. Tolstoy was highly reflective, capable of narrating a complex self-narrative that made sense of his past, his possible futures, and his experiences while he narrated it, and which exhibited an engaged attitude towards his own sense of reasons. Consider the sentiment we might hear from someone in a similar position to Tolstoy, that 'nothing matters, and people who say otherwise are idiots and sheeple'. This uncharitable description is not a defence of philosophical nihilism, but it is a way of defining oneself in relation to others. This agent is offering (perhaps insisting on) a self-conception according to which they are the person who sees the world straight and is honest with themselves and others about its normative state. Depending on how we prefer our theories, this description might establish a practical identity, or constitute the self as an agent, or narrate a self (inclusive of agential capacities and normative relations). Whilst it is based on, and reproduces, a systemically mistaken view of the normative landscape, it can still do the self-defining work that is necessary on various theories of autonomy and at least useful on others. Futility thinkers do not excel in acting as agents, but they can at least successfully stipulate which agent they are, and under what description their agency is to be met.

§1.7.2 Social Conditions for Autonomy

Depressed people are prone to experience the disempowerment that comes with explanatory answers to the *why* question, particularly when explanations trade in pathologizing accounts of depression. The depressed person, it is often assumed, still *really* cares about all the things they used to care about but has lost their motivation to depression. A depressed agent's listlessness is therefore depression at work in the world, rather than the result of their own view of the normative landscape. But insofar as we are more autonomous when we are recognised in the correct way, recognising a depressed agent's

normative outlook extends the relevant recognition. Autonomy is conferred when our interlocutors seek out non-explanatory answers to the *why* question, when they seek answers that pick out how they handle the reasons they handle. Trading in normative outlooks allows us to take a deeper agentic stance towards a depressed person. Rather than rationalising individual actions, or forbearances from action, enquiring after someone's normative outlook allows us to take an agentic stance towards their experiences of value and the patterns in their action, as well as narrowly individuated actions. We don't just confer autonomy onto the actions we understand this way, but onto the pattern of the agent's action (or inaction) and their experience. This is to say, we confer autonomy on the way the agent meets and responds to the existential reality they are faced with.

The depressed agent is still capable of, and engaged in, resolving the world they encounter into a picture of the normative landscape. That they do so with such a restricted colour palate, which prevents them from recognising the normative properties of many facts is an impairment of these capacities, but an impairment expressed only by the faulty exercise of the capacities. The person's agential capacities are what is at work in the world, and for all the errors involved, can be recognised for what they are. As charitable as it seems to extend recognition to the depressed agent as, perhaps, rationally committed to conditioned evaluations, this does not extend autonomy-conferring recognition unless the depressed agent does, in fact, make a conditioned evaluation. To recognise our agential capacities at work means to recognise them as finite, flawed, and prone to various sorts of error. When we are in error, recognition of our agential capacities must involve recognition of our error. Recognition under charitable descriptions fails to be autonomy-conferring if it is not our actual agential capacities that are recognised, it is some hypothetical or idealised capacities.¹⁹ Understanding that normative outlooks are delicate shifting things, and that we will move between different normative outlooks at different times fluidly is valuable for strengthening autonomy in its relational moods.

¹⁹ A qualification: there is some space for charitable or somewhat revisionist recognition to be autonomy-conferring or enhancing when it *capacitates* the agent. We discuss this in §4.2.2.

§1.8 Conclusion

Returning to Killmister's question - what are the ways in which some agent is autonomous?

The futility thinker will fail to respond accurately to their actual reasons. This limits how far their autonomy can be extended in at least some dimensions, however they are not passive in their view of the normative landscape - they actively compose a picture. In each of the passages we considered at the beginning of this chapter, there is a concern with finding where reasons or values lie in the world, and disappointment that it does not seem to be found. These agents gave descriptions of themselves and their location within these barren normative landscapes. They can stipulate how their self is to show up in relation to that landscape. The autonomy that can be built on agential capacities in this state has gouges taken out of it in some, but not all places. The topographical exercise we are engaged in is served by knowing how these impairments transmit up through the edifice and cause depressive or bipolar autonomy to extend unevenly through dimensions, taking on its lumpy, irregular shape.

This account works for typical cases of motivation, but also allows us to track the changes imposed by the affective shifts experienced by people with bipolar or related disorders. Rather than recognising some set of values that the agent holds across time, or is constituted by, trading in changes to one's normative outlook allows us to honour the futility thinkers' testimony that we began with. Rather than possessing some deep selves that they were alienated from, or which was arrested by their depression, they find themselves to be the same people, with new (apparent) moral knowledge. Appreciating depression this way focuses our attention on the agent's perspective on reasons, allowing us to confer autonomy on the agential capacities that are at work in these cases.

Questions such as why our friend missed our appointment can be answered from an agential perspective by recognising that they didn't see getting out of bed as worth it (assuming this is futility rather than accidie). However, we will better appreciate our friend's action, and our relationship with them at this time, if we recognise this as a local expression of their broader normative outlook.

Recognising that our friend is, for now, someone who is blind to a great many reasons recognises them as the kind of agent they are (for now). Insofar as our recognition is necessary for agential capacities to amount to autonomy, recognition of our friend's normative outlook is needed to confer autonomy on

the kind of agent they are. This is why specifying our description more narrowly than 'agent' is needed for mad and neurodivergent agents to flourish *qua* agent. For an agent to flourish as the kind of agent they are, and enjoy the kind of social recognition required to turn capacities into autonomy, the agent must be recognised under a fitting description.

Chapter 2 – Mania and Urgent Agency

2.1 Introduction

Bipolar or related disorders (BoRD) are frequently characterised by manic and hypomanic episodes. Like depressions, these are periods of time in which atypical mood states are expressed, and where these mood states are taken to be disordered or pathological. Mania (and its milder expression, hypomania) presents with a complex behavioural syndrome, tending towards impulsivity, compulsivity, recklessness, obsession, joy, anger, creation, destruction, and excitement. Shoemaker characterises these extremes as involving a loss of “tempering relations” (2015, 134) that hold between our cares, commitments, and impulses. Mania is at least as heterogeneous a phenomenon as depression, too varied to understand simply as a kind of anti-depression, the equal and opposite pole of depressed moods. As highlighted in the introduction, the cultural status of medicalised understandings of madness are such that clinical terminology has become a prominent rallying point around which people gather to share experiences and find common understanding. In deference to this practice, I continue using the clinical terminology and include clinical perspectives. However, the experiences I refer to are not essentially medical, nor does medicalisation offer us privileged insight into these experiences as they make up an agent’s existential reality. Personal autonomy is valuable irrespective of whether we describe our experiences in a clinical register or don’t, and the autobiographical and first-hand phenomenological accounts are at least as valuable in informing us about the existential reality in which bipolar.

Being a less commonplace phenomenon than depression, mania has received significantly less philosophical attention than its (ostensible) opposite pole. For this reason, I will begin with a brief survey of first person and third person observations about mania drawn from autobiographical and clinical perspectives. In this chapter, I will examine the kinds of motivation found during manic episodes and develop an account of the restructuring of agency at work here. If we are to more narrowly specify the description ‘agent’, to better understand the base capacities upon which personal autonomy might be built in the face of different existential realities, then it is necessary that we understand what kind of agent someone is when they are manic. This chapter is an investigation into the restructuring of agency that is distinctive of mania. The previous chapter deployed Tappolet’s sentimentalism and developed

the model to track the changes in our sensitivity to reasons which occur during affective shifts. I argued that the changes alter our normative outlook by changing what facts get marked for our normative attention, and what kinds of evaluations we are able to read these facts in the light of. The salience and valence elements of normative outlooks, outlined in the last chapter, are supplemented in this chapter, where I argue that an additional feature of manic agency marks it out of philosophical attention. This feature indicates that mania restructures agency, rather than simply altering an agent's normative outlook. I characterise this restructuring as a pervasive sense of urgency running through all considerations.

§2.2 will draw on clinical, autobiographical, and phenomenological literature to sketch our initial picture of the phenomenon, loosely individuating various manic moods according to the picture they draw of the normative landscape. §2.3 will draw out the distinctive features of manic agency which we will investigate further for what they reveal about the manic restructuring of agency. In particular, this will be the focus on considerations as demanding, supported by additional autobiographical reports. §2.4 will outline generally what I mean by the restructuring of agency, drawing on Nguyen's account of modular inventories of agency. §2.5 will introduce an analysis of urgency and argue that urgency is the model that best accounts for the distinctive restructuring that occurs during mania. Manic agency, I contend, can be understood as involving a pervasive sense of urgency, as if the mode of agency that is suitable for urgent circumstances has been generalised. §2.6 will consider how such a restructuring of agency shapes our autonomy.

§2.2 Portrait of Mania

Mania is a heterogeneous phenomenon and resists any simplistic unifying account. It is often understood as the polar opposite of depression, but involves much that this description fails to capture. It can be cheerful and optimistic, or bleak and apocalyptic. It can involve high degrees of lability, the seemingly unprompted shifting between moods, but need not. This section will introduce the understanding of moods that I will use, drawn from Radden (2013), but adapted to capture more of the

variety involved in mania. Whilst it is commonplace to refer to mania as a mood, it is better to consider mania as a distinctive way of having moods, (see Fernandez 2014). We can be optimistic, celebratory, or angry in manic or non-manic ways. We will review autobiographical, clinical, and phenomenological accounts in order to build our initial picture of mania. This review is organised loosely into three manic moods according to thematic features of the accounts. This will help us in our subsequent analysis as we draw out common features of experience and motivation. Mania, I contend, restructures agency in a way more pervasive than the nihilistic depressions we examined in the previous chapter. This chapter will offer an account of that restructuring, and its effect on agents' autonomy.

Before examining lived expressions of mania, let us note the most recent clinical description adopted for the purposes of the ICD-11:

A manic episode is an extreme mood state lasting at least one week unless shortened by a treatment intervention characterised by euphoria, irritability, or expansiveness, and by increased activity or a subjective experience of increased energy, accompanied by other characteristic symptoms such as rapid or pressured speech, flight of ideas, increased self-esteem or grandiosity, decreased need for sleep, distractibility, impulsive or reckless behaviour, and rapid changes among different mood states (i.e., mood lability).

(2019, 6A60)

This is the clinical conception of the episodes of mania that feature in the criteria for BoRD diagnoses. This definition involves a number of overlapping and related phenomena which we will characterise in more detail below. Whilst I do not give particular priority to this conception by dint of being the clinical conception in use, it is worth beginning with as the clinical concept is the banner around which people who experience mania have rallied for the purposes of interrogating shared experiences and developing novel, non-clinical, conceptions of mania. The threads visible here will come up throughout the testimonies and reports we will see below.

§2.2.1 The Manic Mood, or Manic Moods

Radden identifies moods with dispositions to particular sets of occurrent emotions which express those moods (2013, 88-9). Emotions have intentional objects and intentional content. (This is 'intentional' in the sense of being about something, rather than of giving an answer to Anscombe's *why* question). A

general fear of dogs is an emotional disposition which finds expression in various emotions. In each instance, the emotion has an object (this dog) and a content (could trample me). She argues that moods are emotional dispositions and that we can speak of moods having intentional objects or quasi-objects.

[disordered moods'] tendency to spawn a host of fully intentional states may explain the way (feature by feature, moment by moment) moods involve intentional descriptions of how the whole world is, creating, in the case of our mood disorders, generalized 'objects' of depression - the world as gloomy and hopeless - and mania - the world as full of promise and possibility.

(Radden 2013, 89)

Moods disclose the world to us, the thought is, by being the disposition expressed by the occurrent emotions which describe any particular part of the world to us at any given moment. Radden distinguishes two moods here, manic and depressed, individuated by the normative picture of the world they offer. As we shall see below, however, mania admits of more variety than just a world brimming with possibility. We might further individuate manic moods into the kind of grandiose optimism Radden considers here, or the bleak agitation that Jamieson remarked on, or the excited and mystical glee with which a manic person might exhaust anyone who will listen by talking obsessively about their most recent fixation. These dispositions to see the world in a particular normative palate, as brimming with promise (optimism), offensive and deserving of ire (anger) or exciting and glorious and worth speaking about at length (celebratory), resolve the world into a normatively rich picture which we can interact with.

A mood is a disposition towards particular emotions, and so overlaps closely, but not precisely, with the normative outlooks we examined in the previous chapter. On Radden's use of the term, moods are individuated according to the picture of the normative world that they paint. Normative outlooks are individuated by the palates they offer us with which we might compose any such picture. Differences between normative outlooks are therefore upstream of differences between moods. The normative description under which the world is disclosed to us will be confined by the range of values we are sensitive to at some time, but the range of values we are sensitive to will often be wider than the range picked out by the emotions we are disposed to experience. We shall return to normative outlooks, composed by patterns of salience-marking and dispositions to valence experience in particular ways, in §2.3.3, but for now I keep the analysis to the level of moods in order to provide a more grounded

picture. I keep these manic moods very loosely individuated, for the purposes of identifying trends in the expression of mania.

§2.2.2 Manic Optimism

Manic moods present a world which “brims with promise and possibility” according to Radden (2013, 84). Attractive as this mood sounds, manic people often make reckless decisions that they eventually regret, being “locked in his irrepressible optimism” (98). The ICD-11 includes, as characteristic of a manic episode, “increased self-esteem or grandiosity, ... [and] impulsive or reckless behaviour” (6A60). Increases in self-esteem and grandiosity both find expression in agent’s assessments of their own abilities, either directly by making more seem achievable, or indirectly, by drawing one’s attention to anything but the risks of one’s plans going wrong. Hence, the manic agent often becomes reckless.

Kay Redfield Jamison is perhaps the most influential scholar of bipolar disorder, owing to her extensive and insightful writing which draws both on her clinical expertise and her own experience with the condition. She reports on one of the numerous spending sprees which nearly bankrupted her, which she engaged in during her own manias. She remarks:

“When I’m high I couldn’t worry about money if I tried. So I don’t. The money will come from somewhere; I am entitled; God will provide”
(1995 74).

Her experience of the world presented a picture in which the things she wanted or thought appropriate were easy, not because she was unaware of her financial limits, but because they failed in some way to seem relevant to her decision-making. In at least some senses, this expresses a particular kind of optimism.

§2.2.3 Celebratory Mania

Whilst this might not be a familiar non-manic mood, I use the term celebratory to pick out those moods in which we are particularly sensitive to the kinds of *appreciation* that values call out for. This appreciation can be expressed by applause, singing, dancing, or making art, but it can also be expressed

by a kind of private mental engagement with the normative properties of something. As we noted in the previous chapter, we can still be engaged with the normative properties of things when we are depressed (although it doesn't seem right to say that one *celebrates* the reasons to rue and decry the evils of the world).

The celebratory thread in mania is found in some of the oldest clinical accounts. Aretaeus observed that manic people “will crown themselves with chaplets and go like victors from a contest” (1837 68-9). This grandiosity is noted in the ICD-11, but rather than one's own grandness having an effect on what seems possible, it has an effect on the kinds of behaviour and appreciation that seem due to one. The grandiose person deserves to be appreciated for their grandness, and the laurel crowns of the Games are as good an expression of that appreciation as any. Psychiatrist Ludwig Binswanger examines another instructive case. He gave his patient, Olga Blum, a copy of Goethe's *Faust* to read. When she completed it, she expressed relief that she lived after Goethe, not because it meant that she could enjoy the novel, but because had it not already been written she felt she would have *had* to write it herself, (Binswanger 1933, 99).²⁰ Binswanger interrogated this remark in order to understand what it reveals about manic grandiosity – why did Blum believe she was equal to the task? However, another reading of her statement is available. This could rather be understood as (perhaps hyperbolic) praise of the novel. We shall return to this case in more detail below, but for now, let us note that Blum seemed preoccupied with the aesthetic greatness of the novel, to such a degree that Binswanger picked out this expression as instructive and worthy of investigation.

Language of celebration is commonplace. Jamieson describes mania as an extreme presentation of exuberance, which she terms “the champagne of moods” (2004, 91). This is a qualified recommendation however, as she also notes that “if exuberance is the Champagne of moods, mania is its cocaine. Mania is exuberance gone amok” (121). Jamieson retains a focus on the celebratory elements of mania, but does not distinguish them strictly (as I have done in an artificial fashion) from the less pleasant aspects. She notes, about living with “the throat of exaltation”, that “Inflammability,

²⁰ Whilst an English edition of this text exists, it is out of print and the two libraries in the UK which list a copy in their holdings are unable to locate them. The copyright of the text has changed hands a number of times since publication and I have not been able to contact the current holder. For now, I rely on the translations included in Martin et al (2019, particularly 82).

however, always lay just the other side of exaltation.” (1995, 122) This impulse to *exalt* can be expressed harmlessly, in the manic tendency to speak at length, at people, about the object of one’s current fixation. These fixations can shift erratically, making manic people hard to follow along with. This shifting of focus is a commonplace, and important for our purposes, feature of mania. In his auto-phenomenological account of mania, Paul Lodge notes the sudden awareness of “an unlimited number of connected objects, or of different ways in which thoughts might be linked together.” (2020) Drawing out more of what we are calling the celebratory element, he describes this as an “encounter ... with something that seems ‘sacred’”, (so worth celebrating if anything is).

§2.2.4 Agitated Mania and Anger

As pleasant as the celebratory strain can sound, it can also be overwhelming. There is an ambivalence that keeps these broadly attractive descriptions from being wholly endearing. Jamieson particularly keeps an eye on the dangers in her choice of language. She is fond of the coinage ‘black manias’ to pick out:

being wildly out of control—physically assaultive, screaming insanely at the top of one’s lungs, running frenetically with no purpose or limit, or impulsively trying to leap from cars— [it] is frightening to others and unspeakably terrifying to oneself. ... [These] black, agitated manias—destroyed things I cherish, pushed to the utter edge people I love, and survived to think I could never recover from the shame. (Jamison 1995 160)

This is a visceral introduction to the agitated sorts of mania that have troubled much of her life. Manic anger can be particularly destructive. As with the excited tendency to exalt, it can flow freely from object to object with little discernment. Travesties in the ancient world will do just as well as the epistemic injustices manic people themselves are prone to experience (perhaps more prone when angry) to be the subject of this anger. To storm and thunder over some fixation serves to express a manic mood just as much as excitedly exalting the contents of one’s most recent Wikipedia rabbit hole. Manic anger is often wrathful and destructive, presenting the world as offensive, and full of injustice at which one ought breathe fire. Describing an episode that seemed to involve a great deal of manic anger, Terri Cheney writes:

“I woke up early the next morning full of energy, eager to go, and irritable as a drenched cat. Everything Rick did annoyed me, from the way he tapped each side of his soft-boiled egg six times precisely, to the

way he said “love you,” without the “I” ... I hated the world, I hated myself, and dying sounded just fine to me: all classic symptoms of depression. But—and it was a crucial but—I could still move. Not only could I move, I had to move. I was full of restless, undissipated energy that had no place to go, making me want to strike out and break something, preferably something that would crash and tinkle into a thousand satisfying tiny pieces.”
(Cheney 2008, 183-4)²¹

There was no particular thing about which Cheney was angry. This anger could just as well be about the world, rather than any particular event. It is not surprising that this is one of the more unpleasant turns mania can take and is one of the more frightening for one’s friends, family, and even oneself.

These examples of mania paint a mixed, frequently unpleasant, picture of an emotional and motivational state. Per the Sentimentalist thesis, we should be able to say something about how these emotional states are representing reasons to the agent. In broad strokes, we could recognise a few different manic moods here, offering distinct pictures of the normative landscape. The optimistic or celebratory mood presents the world, including the future, in a favourable light, perhaps even as glorious. The agent's attention is fixed on these glorious features of the world, and on the reasons to appreciate or remark upon them. Manic anger casts the whole world in terms that are personally and individually insulting to the agent, and anger at the mistreatment of others, or of ancient historical events, is indistinguishable from severe offences against oneself. We can recognise broad features of the various normative outlooks beneath each of these moods in the descriptions given. The general agitation that Jamison remarks on portrays the world as uncomfortable or unaccommodating, marking for salience all the wrong things, perhaps valencing things that would otherwise seem neutral as negative.

²¹ Chaney characterises this as a mixed affective episode (see ICD-11, 6A60.9), a term that she reports encountering for the first time during this period. The only elements of her experience that she notes as indicating that this was not a manic episode, however, is the negative (unpleasant, unhappy) valence to her affect. For this reason, I am more inclined to consider this a simple manic or hypomanic episode. Whilst Cheney reports that this felt starkly different to her usual manias, it seems very plausible from her account that this was a manic episode with what was for her an unusual expression. If it was a mixed episode, however, the manic elements remain instructive.

§2.3 Manic Agency

The manic moods outlined above involve rich pictures of the normative world. In this section we will develop a more specific picture of the agentive changes that mania occasions. Having drawn our general picture, I will below analyse these accounts in further detail, adding neighbouring accounts where required, to demonstrate the distinctive motivational character of the manias we are considering. The distinctive feature of manic motivational states, as we shall see, seems to be *demandingness*. We will consider how salience marking and valencing might contribute to this change, but conclude that these alone are not sufficient to explain the change. Mania, therefore, occasions a more pervasive restructuring of agency.

§2.3.1 Narrow Thought-Action Gap

Returning to Jamison's autobiographical examples, she reported on one of numerous spending sprees, some of which nearly bankrupted her, which she engaged in while manic. Extending the quoted report above, she tells us:

"I couldn't worry about money if I tried. So I don't. The money will come from somewhere; I am entitled; God will provide. ...So I bought twelve snakebite kits, with a sense of urgency and importance. ... During one spree in London I spent several hundred pounds on books having titles or covers that somehow caught my fancy: books on the natural history of the mole, twenty sundry Penguin books because I thought it could be nice if the penguins could form a colony. Once I think I shoplifted a blouse because I could not wait a minute longer for the woman-with-molasses feet in front of me in line. or maybe I just thought about shoplifting, I don't remember, I was totally confused."
(1995 74).

Optimism does not seem adequate to capture this. Worries about money were not absent, but impotent. An opportunity presented itself to produce the (rather wonderful) practical pun involving paperbacks huddled for warmth on the coffee table. It was an opportunity so good that, from Jamison's point of view, what else was she to do?

The tendency to act without much concern for consequences is a commonplace and commonly remarked on feature of manic behaviour. This example speaks to the wider trend noted in the ICD-11's symptom list, including "distractibility, impulsive or reckless behaviour" (6A60). The gap between the reason to do something becoming apparent and the agent's action is very narrow. There is not space here for deliberation. This holds as true for what gets our attention as for what we do. Manic agents are distracted by the new thought just as easily as they are motivated to action by a new thought. That there isn't space for deliberation between appearance of reason and action produces the recklessness that the ICD-11 notes. Indeed, this is largely why agitated and angry manic moods are terrifying. As agitated as one feels, as much as the world prickles and seems offensive, if there is a wafer-thin gap between apparent reasons and actions, the agent can be left fearful that they will snap and lash out at any moment. This disposition is expressed in practical behaviour. If we are to develop a more rounded account of manic agency, we must ask what phenomenology underlies this disposition? What does it feel like from the inside?

§2.3.2 Mania and Demandingness

Consider again Olga Blum's testimony: if she had lived before Goethe, she would have *had* to write Faust. Binswanger and his followers interrogate this remark for what it tells us about manic grandiosity. They ask why Blum considered herself up to the task, (Binswanger 1933, 99, Martin et al 2019, 82). We noted that there was another interpretation. At the most literal level, Blum is stating that had the novel not been written, she would have been compelled to write it. I understand this as making a normative, rather than motivational claim. She suggests that the world *should not be suffered to continue* without so great a work of art in it. The novel is of such artistic value that, irrespective of how difficult it might be for her to produce it, the good to be won makes that worthwhile. This is not an arbitrary interpretation. It is not a strikingly odd expression translated into English, and the expression 'if it weren't for (you/it), we'd have to invent (you/it)' is a common idiom in Hungarian.²² How difficult it would be for her to write is not itself relevant to whether she should do it, as long as the stakes are sufficiently high. (Consider, is there any possible degree to which averting climate disaster could be

²² I am indebted to Zsuzsanna Chappell for edifying conversations about this case, and whose familiarity with the Hungarian idiom provided the initial insight. A student tells me the same idiom exists in German.

difficult that would make it not worth doing?).²³ If we understand Blum to simply be referring to what would be her reasons in a world where Goethe had not yet lived, we can see the kinds of reasons she took seriously. The world is one in which these great artworks could exist, and if we fail to seize this opportunity, we doom the world to persist without them, which we must not. If Blum's comments help us characterise general features of mania, understood this way, they suggest that mania involves an experience of reasons as *demanding*.

The language of demandingness is all over the literature. One of the two explanations Jamieson offered for her shoplifting was that she could not wait for a slow-moving other shopper. In another work, she shares the account of an anonymous person affected by mania during the early 20th Century:

“All the problems of the universe came crowding into my mind, demanding instant discussion and solution ... I felt like a person driving a wild horse with a weak rein, who dares not use force, but lets him run his course, following the line of least resistance. Mad impulses would rush through my brain, carrying me first in one direction and then another.”

(in Jamison, 2004, 241-2)

This agent is moved by reasons, but without the opportunity to weigh them. She was not able to steer her thoughts, and the variety of subjects she was drawn to was similarly beyond her control. It is not up to her what she dedicates her attention and practical efforts to. Where neurotypical agents are more used to having various objects of attention and being able to sift between them to decide what to focus on, this agent was subject to the demands of “mental telepathy, hypnotism, wireless telegraphy, Christian Science, women’s rights, and all the problems of medical science” (ibid).

Those contending with manic anger have to work constantly to manage the relationship between their actions and the picture of the world, as deserving everywhere to be acted on, waxed lyrical about, or be lashed out at. In an article for *Speaking Bipolar: Keep Fighting*, an anonymous author recounts their experience with manic anger:

Right now I could break something. Whether it be the pen in my hand or the mirror across the room, something needs to be destroyed. Bipolar anger says it must be done. ... Bipolar rage ... [is] like being a

²³ There is not.

hungry lion and someone has just tried to drag your last and only food source away. There will be blood.
(Speaking Bipolar)

Again, we see a shift to talk of what is demanded. This agent sees their anger as making demands on them rather than presenting reasons. Anger, and to an extent bipolar disorder itself, is reified and made into another entity to which these demands can be ascribed. The agent's own role as the person who is angry, and whose anger presents reasons to decry or to remonstrate, is omitted. The agent seems to be along for the ride.

The recurring theme of demandingness is, I contend, crucial to understanding the way mania restructures agency. This is an experience of value in which the agent finds the normative properties of the world to set rigid standards that they are measured against. The normative properties that made some response fitting now makes it an expectation. I am fond of Carrie Fisher's summary: "when you're manic, every impulse feels like an edict from the Vatican" (2009, 128). The suggestion of divinity about the manic person's projects, and the not towards papal infallibility, tie together the demandingness and cosmic profundity of the manic experience of value. Manic agents act without much deliberation, but it is not clear that this is any kind of failure by their own lights. Indeed, it seems that the considerations they respond to are ones that call out for immediate, swift, or decisive response.

§2.3.3 Agentive Capacities and Agentive Structure

Is our account of normative outlooks sufficient to capture this demanding quality of mania? Let us consider how the manic agent's emotions are marking facts as salient and valencing those facts. The practical pun created by Jamieson's crowd of paperbacks, huddling for warmth on the coffee table, is really quite delightful. The reasons to create this pun, the aesthetic properties of such an arrangement, were marked salient for her. The writer for *Speaking Bipolar* has his attention fixed on the reasons to be angry and wrathful. These groups of reasons consumed the agents' attention. Some normative properties are marked for salience, but the salience is extreme and emphatic. Manic moods seem to disclose the world with exclamation marks over all its normative properties. Normative properties are loud, in comparison to what is normal.

However, mania does not make all reasons seem weightier in unison. Some normative properties go silent, their corresponding reasons becoming featherlight. Jamison simply could not make herself care about her overspending. That she remarks on being unable to care about it indicates that she was at least aware at the time that her income being finite did in fact speak against spending very large amounts. This reason was available to her, she was capable of recognising it as a reason, but it could not attain any significant weight in her deliberation. This is unlike the case in futility thinking, where some families of reasons simply fail to show up as reasons at all. The futility thinker rejects the claim that it is really worthwhile to engage in various self-care practices. Jamison, on the other hand, was able to recognise the normative salience and valence of this consideration, but for some reason it did not move ground in her deliberation.

There is a further complication regarding salience. The unnamed historical author found not only that some reasons were salient and some were not, but that everything that was marked for salience came in a mad rush. Just as she fixed her normative attention on one thing, something else demanded it. Paul Lodge has recently explored this theme in some detail, (2020). The central theme in his auto-phenomenology is an awareness of “an unlimited number of connected objects, or of different ways in which thoughts might be linked together.” The usual salience marking that we rely on to draw out the parts of the world that are worth deliberating on or interacting with goes into overdrive. The normative properties of things that we are sensitive to become stark and loud, but more importantly, they come thick and fast and the agent might struggle to keep up, and become overwhelmed. This is the distinctive pattern of salience marking that is at work when our emotions report on the normative landscape during manic episodes.

Neither do we have a wholly straightforward story about valence. There seems to be a step difference between the way neurotypical onlookers might appreciate the aesthetic value of Jamieson’s pun, and the way these reasons are appreciated by her at the time. The *to-be-doneness* is starker for some reason. Each particular normative property that an agent detects is regarded as bearing great profundity. “At the heart of the experience” Lodge remarks (ibid) “is the feeling of successful insight.” Mania, taken to be revelatory by people who experience it, brings a sense that one’s normative outlook has been brought to some more profound level, and one now sees the world more clearly than before,

(even if one is suspicious of that sense). The meaning in the world is greater, and more important. Both the frenetic intensity and the flightiness of mania are accounted for by the agents' always encountering very powerful reasons to do something, which requires all of their efforts and justifies dropping everything to attend to immediately. Recall that the valences we apply to our emotions are constrained by the kinds of intentional state facilitated by the existential feelings which underly our present mood. Mania does not only involve a busier salience pattern than non-manic celebratory moods, it also involves a sense of profundity or of 'the sacred'. This is a difference not in the particular valences (positive or negative) that we apply to our feelings, but a difference in the *texture* of those valences. If the existential feelings theory is correct, then we might expect that manic moods are underpinned by foundational existential feelings which facilitate valences with these profound, mystical, or sacred textures.

The loud normative properties disclosed by manic moods are more important properties, or properties which correspond to weightier reasons to act in certain ways, or to hold certain attitudes. Thus, the thing to do is easily settled by the thought that has just occurred to one. The valences are qualitatively different, prone as they are to a sense of profundity and a leap to the cosmic. An analysis in terms of salience and valence alone was sufficient in Chapter 1 to account for futility thinking, however we noted that more is involved in our emotional experience of value than only salience and valence. So far, this is inadequate to explain demandingness. The manic agent has a noisy and busy salience pattern, and valences their emotions in ways that tend towards profundity and mystical feeling. What is not clear is why either of these features cause the manic experience of value to involve *demandingness* rather than simple (perhaps stark) fittingness. Mania, therefore, is not accounted for only by a shift in normative outlook, the way that nihilistic depressions were. Manic agents are reasons-responsive in an uncommon way, taking an appropriate response to reasons to be more demanding. This is a more pervasive restructuring of agency. The remainder of this chapter will develop an account of this pervasive restructuring.

§2.4 Modular Agency and Agential Structures

The purpose of investigating manic agency is to provide us with a more specific description of what is involved in being this kind of agent. As outlined in the introduction, I take autonomy as an ideal to involve the excellent functioning as an agent. However, the descriptor ‘agent’ is too broad. We are all different kinds of agents at different times, and so the relevant ideal of autonomy we might hope to aim for will change subtly, or not so subtly, as the kind of agent we are shifts. The utility of an account of manic agency in particular is that it specifies the kind of agent at issue here and so can inform us of what excellent functioning as that kind of agent might look like. In short, we need to know what kind of agent we are when manic to know what the ideal of autonomy looks like local to that kind of agency. Before arguing for any particular local model of agency, I will briefly outline how I understand structures of agency to vary.

A good place to start is with Nguyen’s (2019, 2020) modular view of agency. Nguyen argues that when we play games we submerge ourselves in *layers of agency*, which involve distinctive ways of recognising and responding to reasons. When playing games, we recognise otherwise trivial or normatively neutral facts as reasons. While playing *Risk*, an opportunity to betray your trust might arise and whilst I would normally never dream of such a thing, this opportunity does present a good reason to do so. The reasons we respond to during a game of *Risk* are *self-effacing*, in the sense that we (presumably) do not actually value winning the game more than we value the diversion and bonding that boardgame night offers, (2020, 54).²⁴ We respond to self-effacing ends by not aiming directly at them, rather by confecting other ends so that we might ‘sneak up on’ them. This is a different way of handling practical reasons from that which we use for more typical non-self-effacing ends. Self-effacing ends are commonplace, we interact with them if not daily, then at least weekly. We can nimbly shift between modes of agency which handle reasons in the ways required by regular non-self-effacing ends, and the ways we handle self-effacing ends. This dexterity is a typical part of most people’s agency and does not

²⁴ Whilst the reasons we adopt in games are self-effacing, they are still real practical reasons with genuine normative force. The goods presumed in the rules of *Risk* are already goods I have committed myself to by participating in the game, and so I would be practically irrational to reject the reasons illuminated by those goods. In my terms, we submerge ourselves in this layer of agency, in part, by adopting a restricted normative outlook given to us by the rules of the game.

require a specified description of agency that travels any distance from commonplace neurotypical types. These shifts into artificial normative outlooks, and into distinct ways of *handling* practical reasons, mark out the different layers or modules of agency we can adopt.

This is broadly how I understand the restructuring of agency, however our case differs from Nguyen's on the following point. On Nguyen's theory, all agency is highly modular, and each of these structures is a tile in the larger mosaic of our agency (an 'inventory of agential modes' as he terms it, 82). Crucial to this theory is our capacity to submerge ourselves in a layer of agency. When we play *Risk* we adopt this narrowed normative outlook and new norms for handling reasons for the purposes of playing the game, but we retain contact with our background agency. We remain sensitive to reasons to stop playing if people aren't having fun, conversation is bubbling along nicely on its own, or a fire breaks out. On any of these eventualities, we know to withdraw from the layer of agency we have adopted (57). This is the description he offers to layers of agency, which are only loosely individuated from our overall agency. This is where mania differs. None of our testimonies or clinical observations indicate *submersion*. Manic agents cannot readily withdraw from their frenetic layer of agency as and when required. The affective shifts imposed by bipolar and related disorders do not offer narrow normative outlooks we can adopt like spectacles, for the purposes of playing a game. They are changes in the kinds of response agents take normative properties to call out for. The modules of agency that Nguyen describes, structures of agency which we can store in an inventory and fluently deploy when needed, are small-scale models of the kind of far-reaching restructuring of agency at work during manic episodes. The key similarity is that mania does not only introduce a distinct normative outlook, as nihilistic depressions might, but also changes the way agents *handle* the reasons they see. The dissimilarity is that the whole of one's agency goes with this change. The manic agent has a normative outlook that tracks reasons, in a loud and perhaps garish palate, but what constitutes an appropriate response to the reasons tracked is changed. This is the restructuring of agency.

Nguyen's account is instructive in identifying the ways our agency (or agencies) can be changed by circumstances, efforts of our own, and practice. The kinds of change he has in mind are narrow and local and specialised to the tasks in hand (boardgame night) or developing the kinds of inventories that can make us more dexterous agents overall, (drawing on a library of agencies). Removing the layering

and submersion element allows us to apply this account to the broad restructuring going on in manic episodes.

§2.5 Manic Agency as Urgent Agency

In this section I will argue that we can understand manic agency as restructured in a way that involves a pervasive sense of urgency. The kinds of response to reasons that mark out this restructuring are the responses distinctive of agency in urgent circumstances. This is a mode that all effective agents can adopt when needed, but seems to be a key feature of agency as restructured by mania.

§2.5.1 Urgency

Fighting fires is urgent. Hippocrates remarks that whilst medicine is years in the learning, there is usually only a moment (*kàiros*) available for it to be effective (2015, 798). Artistic projects are sometimes like this. Michelangelo said that he saw David in the marble and was compelled to free him. Very plausibly, freeing David became urgent as soon as Michelangelo saw him. If Michelangelo takes too long, it isn't David that emerges from the marble, but some lesser creation. Either because the idea became overgrown, or because the moment in the unfolding artistic context is missed.

Urgency is a normative concept. We invite criticism if we fail to respond to urgent considerations with appropriate haste or efficiency. I take urgency to pre-empt further deliberation. Whilst usually this is because we do not have time to deliberate, it can also be because we lack the logical space for (permissible) deliberation, for instance, when we are in situations of *moral* urgency. Some injustices are urgent not because they can be rectified only before some particular time, but because *immediate* response (be it apology, comfort, or even just recognition) is what the urgent reasons call out for. Some balance of reasons being decisive does not make matters urgent. We can have best all-things-considered reasons to ϕ rather than ψ and still enjoy wide latitude to deliberate. In urgent circumstances, however, further deliberation is not superogatory, but *inappropriate*. Whichever

reasons are decisive also pre-empt other considerations. Our reasons not to soak rare books do not lose their normative force when fighting fires in a library, but we should not get so far in our deliberation as to be weighing them up. To do so is to indulge in "one thought too many" (as Williams has it, 1981, 18) Hippocrates highlights the need to make the right decisions without the luxury of deliberation. An adequate response to our urgent reasons is to close deliberation. Anything short of closing deliberation is an inadequate response to these reasons.

Note that this is not the same kind of pre-emption that Raz discusses, (1986, 52). On his view, legal authorities pre-empt deliberation by 'scooping up' and replacing our practical reasons. My reasons to contribute to the funding of the health service are not for me to weigh up, they are (Raz argues) properly incorporated into my legal responsibility to pay tax, by a legitimate political authority. No such substitution occurs in urgency. This is a difference not in the reasons to which we respond, but in what counts as an adequate response. To respond to a morally urgent injustice with uncertainty and equivocation is to fail to respond adequately, even if one later comes to recognise one's mistake. To fail to respond to urgent circumstances with decisiveness, fleet-footedness, or efficiency, is to fail by the normative standard set by urgency.

The analogy between urgent situations and the way manic agents handle their reasons is revealed when we reflect on the demandingness of the reasons that are apparent to manic people, and in the impulsiveness with which they act. Mania presents an opportunity to achieve something valuable, but for which there is little space to deliberate, either because there is no time, or more directly, because there is little logical space to permissibly deliberate before acting on practically (morally, aesthetically) urgent reasons.

§2.5.2 Demandingness and Space for Deliberation

Games involve structures of agency which we adopt and which see us handle reasons in distinctive ways. In games, our ends are self-effacing, so the reasons we respond to require that we not think about our final ends and trust in the framework we have adopted in order to approach them elliptically. To the extent that manic moods restructure agency, they change the way manic agents respond to

reasons. The change effected in mania, I argue, is a change after the model of urgency. The manic agent responds to their reasons as if they were in urgent circumstances. Two distinguishable thoughts are closely wound together in the passages which convey this sense of urgency – one of being practically without control of one’s mind, actions, or situation, and another of being *normatively* without control. Like the agent in urgent circumstances, the manic agent does not find that they enjoy wide latitude to deliberate.

The anonymous author that Jamieson quoted remarked on the demands that each thought placed on her. She opted for the metaphor of trying to control untamed horses with inadequate tack in order to convey her experience of her own mind. Comparing this with other descriptions is instructive. The clinical diagnostics place great emphasis on impulsivity in mania, as does Jamieson who includes “running frenetically with no purpose or limit, or impulsively trying to leap from cars” as among the more disturbing of her own behaviours while manic (1995 160). These descriptions put together emphasise the lack of control that manic agents experience – with a wafer thin gap between idea and action, it is natural to become wary of one’s own mind, or to become afraid as one does not know what one will do next. This is the feeling of lacking practical control of one’s own mind.

Yet, we can compare the anonymous author’s description of the demands placed on her with two further descriptions. Each of the questions the anonymous author faced, telepathy, women's suffrage, Christian science, were not only worthy of some attention, but placed demands on her. Fisher tells us that “when you’re manic, every impulse feels like an edict from the Vatican” (2009, 128). The use of religious language establishes certain stakes. Divine authority and papal infallibility (both usually understood to be exercised over normative questions) are the metaphors she uses to describe the way reasons seemed to her during her manias. Olga Blum’s own literary abilities were irrelevant to what she should have done, all things considered, were she to have lived before Goethe. These descriptions do not focus on the practical effect the agent feels they can have on their situation, mental state, or actions, but on the reasons they have to act. These reports of manic experience are also instances of normative testimony. The agents face normative situations in which everything is a demand, and they lack either the standing or the space to deliberate further. This is what it means to be normatively out of control.

Usually, reasons require things of us only when they stand in some defeasing relation to other reasons, which merely suggest things to us. The change here is in the handling of reasons. The weighing metaphor, which is widespread, ceases to apply as elegantly as it usually does. During a manic episode, the agent sees reasons arrive which are to be acted on at once. This follows the pattern of urgency. If the firefighter or physician does not act now, there will be no point acting at all. There is no space (in this context, time) for deliberation. Olga Blum had time, but the values available were such that there was no need for deliberation. The novel was so good that it simply had to be in the world, and so deliberating further would simply be putting off what was to be done. These are cases where the agent lacks a kind of normative space to deliberate further about what to do. The response that these reasons seem to call for precludes further deliberation. There are some contexts in which it might be better to speak of mania creating the impression in an agent that they lack the standing to weigh their reasons further, but we shall explore this more in §2.5.3. For now, we have the beginnings of the restructured agency that is distinctive of mania.

Given this diminished space to deliberate, how is the manic agent taking their reasons? Usually, reasons are ubiquitous and come cheap. Any fact can reveal a practical reason in the light of some appropriate good (again, Alvarez 2010). My students have reasons to bring me chocolates every day in class, as it *might* make me happy, and reasons not to, as it *might* ruin my teeth, (weak-willed as I am, it almost certainly would). Reasons are usually only contributory. One might have any number of reasons to act in certain ways, but it is not until some balance of reasons obtains that it is the case that one ought (all things considered) act in that way. Reasons speak in favour of (or against) some action or attitude but do not speak decisively. Normally, we can weigh and dismiss reasons, and frequently reasons have no claim on even being entertained. My students probably never consider the reasons to bring me chocolates in class, but they do not invite any moral or rational criticism. They are thoroughly ignorable reasons.

Sam Shpall (2014) offers a helpful analysis of the different kinds of normative considerations we routinely trade in. Two features which reasons lack, but which exist in other normative considerations

such as requirements and commitments, are *strictness* and *decisiveness*.²⁵ Considerations that are strict are those considerations that we cannot take or leave. Shpall suggests that reasons are non-strict whereas commitments are strict (158), however we can easily imagine some reasons that we might consider strict, such as the reasons to be invested in our own and other people's flourishing. They might not be ones we must deploy in all our deliberation at all times, but we cannot opt out of morality as and when we feel like it. These reasons are strict in that they lay claim to at least some of our attention. Decisiveness is distinguished from *pro tanto* status, (154). Reasons which are strict might not be decisive: I don't get to dismiss the reasons to do with my own or my friend's personal flourishing however I can find that they are outweighed by other considerations. Considerations that are both strict and decisive Shpall refers to as requirements. Requirements of rationality, or of morality, tell us what we should do all-things-considered. No normative considerations are powerful enough (even in coalition) to overcome our requirements, and we don't get to dismiss our requirements. You might have some reason to believe that your preferred candidate won an election: it makes you feel righteous and justified in your policy preferences to imagine that the (silent) majority backs you up. Rational requirements, however, demand that you believe that the opponent won this time, as that is what all available evidence indicates. To believe that your candidate won, then, is to fail to be rational. (You might even be rationally *committed* to your irrational belief, however you still violate your requirements, and are in defiance of your decisive reasons.) Requirements are not escapable. Whilst a friend might release me from a promise, they cannot release me from the strict requirement that reasons grounded in their flourishing have a certain weight in my deliberation. What we are rationally, or morally, required to do is inflexible in this way.

This is the kind of response reasons seem to call out for in mania. This changes not only what reasons we discern, but what reasons look like when we discern them. The reasons we discern in this state seem to call out for a different kind of response, and we seem to lack space to deliberate.

Consequently, we take the reasons we encounter as if they were requirements. They seem to us to be

²⁵ Shpall seems to take reasons, commitments, and requirements to be different kinds of consideration, however I am inclined to take commitments and requirements to either be sets of reasons in conference together, or reasons in some particular context or situation that changes how they are to be handled. Nothing in my argument turns on this distinction, which is largely a terminological difference. I will refer to all these considerations as reasons, specifying whether they are strict or decisive or otherwise. When I refer to requirements I am referring to (clusters of) reasons that are (cumulatively) both strict and decisive, and not some *sui generis* normative consideration.

decisive, and often strict, no longer merely contributory. By narrowing the space for deliberation, mania restructures agency after the fashion of urgency. The agent, like the firefighter, has very limited space to weigh up the reasons they detect, to indulge in this deliberation would be to fail by these reasons' demanding standards. Jamieson could still recognise that her finances gave her reasons to stop spending, but these reasons could not win out in her deliberation, because other reasons appeared to have the normativity of requirements to her. Olga Blum didn't need to consider her abilities as a writer, as no facts about that matter would change what she should do. Urgency obtains when our reasons are of this kind. Some urgent situations do not have one particular thing we must do or outcome we must achieve (such as extinguishing a fire) but have a thing we must avoid (the patient dying) which establishes the decisiveness standard. If the reasons that are creating urgency are merely pro-tanto rather than strict, then we can escape urgency by dismissing those reasons. Most situations that are urgent are so either because we cannot dismiss those reasons (they are strict) or because we are already committed to them in some sense.²⁶

We should understand manic agents to handle their reasons as if they are strict and decisive. This structure of agency is reminiscent of urgency in this regard. Before we conclude by considering what kind of agency this is and what its excellent form might look like, let us briefly consider two variations in the structure I have outlined here.

§2.5.3 Hyperagency and Hypoagency

I will briefly outline the degrees to which this structure can create a sense of hyperagency, or hypoagency, as where on this spectrum a manic structure of agency falls will have consequences for the exercise of capacities and excellence in their exercise. By hyperagency, I refer to a sense that one's own agency is more significant than that of other people, that one has been picked out for special attention by the normative properties of the world. By hypoagency, I refer to the inverse, the sense that one has no control or standing to weigh in on reasons and must simply go as one is directed. These are not

²⁶ We can still escape urgency by revising our commitments. This is sometimes a rational thing to do, and is sometimes even an exercise of our autonomy (Radden 1994). Being excessively ready to revise one's commitments, however, is injurious to our autonomy in one of a number of dimensions depending on your theory. See §5.5.1 for a discussion on the threshold of rationality here.

distinct structures of manic agency as one's handling of reasons is significantly similar, however they lend themselves to different kinds of relationship to the self, and so will have consequences for one's autonomy.²⁷

Manic moods involve hyperagency to the extent that they present a view of the normative landscape on which normative properties and their associated reasons refer specifically to us in some way. That some action would make another person happy is a justifying reason any agent could deploy in their practical reasoning - however we usually grant that some reasons are reasons for some agents in particular, and so fail to speak to other agents in the same way. We all have reasons to ensure that a child is nurtured and provided for, but not the same reasons that the child's parents have. Any of us might come to owe an apology, but once we do, the reasons to apologise are not background facts of the moral universe, they seek us, in particular, out.

A distinctive feature of manic moods is that the normative properties of the world (seem to) disproportionately refer to us in a way that they do not when we are not manic. The world disclosed by manic anger is brimming with offence aimed at the manic agent in particular. Manic anger casts the whole world in terms that are personally and individually insulting to the agent, and anger at the mistreatment of others, or of ancient historical events, are indistinguishable from the severe epistemic injustices one is more prone to while manic. At the other extreme, when Jamison could not worry about money, she concluded that "the money would come from somewhere. I am entitled; God will provide." (1995 101). The manic agent seems to be entitled because normative properties of the world (apparently) fall about for the agent's own convenience. The reasons that pick the agent out are experienced as demanding and decisive, narrowing the space for deliberation. But the attention these

²⁷ Petrolini (2021) has used these same terms to refer to a sense that "events that are completely unrelated to her actions (or thoughts) fall under her own responsibility and therefore experience unbearable guilt as a result. ... In these cases, subjects attribute to themselves a greater degree of agency and control than they actually possess." (80-1). And in the inverse: "[A] person may be unsure of whether she initiated an action that others attribute to her, or she might deny having done so despite evidence to the contrary," which she calls hypoagency. In both cases, this altered scope of agency is backwards looking and picked out with reference to the self-reactive attitudes the agent has which are unfitting. It is also limited to the scope of done actions, rather than unfulfilled intentions or the motivations that never make it to being acted on. My use of the terms encompasses significantly more.

reasons pay to the manic agent marks them out as special. Grandiosity might play a role in this way of understanding the world. When you are the most excellent specimen there is, there are plausibly stronger reasons that you should do certain (excellent) things. Indeed, grandiosity might be quite epistemically innocent, if it is the only way of making sense of an experience of value according to which one seems to have been picked out by the world's normative properties for special attention.²⁸

Hypoagency is the style that obtains if the agent feels that they are merely being commanded this way and that and that their own perspective is irrelevant. This is the sense in which the space is closed not because the stakes are high, but because the agent doesn't have standing to weigh in. The agent is not the arbiter of what they do, but the vessel by which reasons have their way with the world. Artists sometimes talk about being absent or uninvolved in their artistic works, ('it writes itself', Michelangelo 'freed' David from the marble block). There is an obvious sense in which this is false, but the artist might be better able to enjoy their own work as art by detaching from their agency to some degree. Manic hypoagency generalises this approach. In the Speaking Bipolar excerpt, we were told that: "Something must be destroyed. Bipolar anger" and not the agent, "says it must be done." Yet it remains for the agent to do it. Manic agents can feel that they do not have normative control when demanding reasons appear and overwhelm them, and they cease to see themselves as more than taking instruction from manic ideas. The agent is, like the artist imagines themselves, the vessel for the universe, in its Glory, its Fury, or its sheer Urgency, pouring out into the tangible space of action.

Interestingly, hypoagency needn't preclude hyperagency. The manic individual might be perfectly content to be God's instrument for acting on the world, understanding themselves to be God's intended audience for the grand drama of creation. The agent does not need to exercise their own judgement if reasons announce themselves and need only be acknowledged (by acting accordingly, or just appreciating value in the right way). The result, on this picture, is still that the goods and bads of the universe will orbit the manic agent in the relevant pattern.

²⁸ Bortolotti argues that a belief, even a delusional belief, is epistemically innocent insofar as it confers epistemic benefits (like intelligibility of one's situation) that would otherwise be inaccessible, (2015, 2020).

Sometimes manic structures of agency will occasion hyperagency, sometimes hypoagency, and sometimes both, all to different degrees. One might be hyperagentially angry, and continue to exercise strict agential control over what it is appropriate to be angry about. Indeed, shifting around between various outrages to be angry about is the familiar manic flight of ideas taken in a particular key. Conversely, one might believe oneself to be the beating heart of the universe, and might do whatever the Universe tells one to do, drawing only a wafer-thin distinction between self and world.

§2.6 What Kind of Agency

The first section of this chapter introduced us in broad strokes to the kinds of experiences that are involved in mania, and the second characterised the specific experiences of value and motivation that manic people report. The third section outlined how I understand structures of agency to be distinguished. The fourth section furnished us with an account of how agency is restructured in mania. A manic agent's emotional reasons-responsive system begins to mark salience in a much more erratic way, and valence marking creates the impression that the normative properties of the world are louder or more profound than usual. Reasons appear to the agent as calling for a different sort of handling, seeming strict and decisive where they would usually be recognised as *pro-tanto* and, in at least some cases, optional. The manic agent operates in a narrowed space of deliberation, responding to apparent requirements coming from all directions. This might be an exciting prospect, as their profound sense of insight leaves them feeling like they finally know what is to be done, believed, or appreciated in the world. Or, mirroring the feeling that they cannot control their situation or thoughts, they might find themselves subject to an additional arbitrary force in the reasons they are discerning. There are various steps one can take to manage these shifts in one's agency, and to help return one to a more familiar and accommodating structure of agency. Some of these are self-care practices, some of these are pharmacological interventions. There is, however, a limit to how much can be done to mitigate this restructuring. I focus on mania, depression, and bipolar disorder because they involve broadly episodic shifts in our capacities to recognise and respond to reasons and in the overall structure of our agency. These shifts make up part of the existential reality for people living with bipolar or related disorders largely because they can be managed but not wholly eliminated, and managing them takes continual

effort. The manic restructuring of agency, therefore, is not something that can be eliminated. The kind of agency enjoyed by people with BoRD is an agency that involves episodes of this restructuring. Thus, to understand the ideal of autonomy local to this vicinity, we need to understand what the excellent functioning of this kind of agency looks like.

§2.6.1 Manic Urgency

Part of the existential reality facing people with bipolar or related disorders is that their agency will be restructured in this way from time to time. What, then, does the excellent exercise of this structure of agency look like? The first conclusion we must come to is that manic people will frequently be wrong about the requirements they face. Taking most or every salient reason as a requirement creates a warped picture of the normative landscape (although less warped than that of futility thinking) and will lead one to do the wrong thing sometimes. However, it does not need to lead one that far astray. Whilst each apparent requirement might, in fact, only be a contributory reason, the manic agent still recognises these considerations as speaking for, or against, the acts, thoughts, or feelings that they do in fact speak for or against. What is lost is the kind of ‘tempering’ that Shoemaker holds is provided by clusters of cares and commitments, (2015, 135). The agent sees single features of the normative landscape in meticulous detail, but will struggle in tasks that require seeing the balance. The totemic importance of whichever reason seems so demanding in this moment is enough to eclipse the normative force of other reasons the agent would normally weigh heavily, and so their judgement will be both unusual for them, and frequently destructive to their long-term projects and relationships.

However, there is still a space for this kind of agency to get the agent’s reasons right. It is a matter of cultivating self-care practices and habits that help one to focus manic attention on the right things. The creative arts are an excellent example, as art works can be let out of one’s head into their first draft, and then left fallow for a period without deteriorating the way that, for instance, relationships do. Jamieson has given specific attention to the relationship between manic depressive illness and an artistic temperament, (1994). The shifting between extremes of feeling, the fluency with which manic depressives cross various kinds of border, and the importance of these qualities in art are, she suggests, a conducive match. There are domains of life in which the kind of urgency I highlight as a feature of the manic structure of agency makes an agent more likely to meet their mark. Where these domains are

and how to cultivate reliable skill for them is the subject of Chapter 3. Doing so without giving into just whatever thought or feeling mania presents requires a finely balanced self-care regime, which is explored in Chapters 4 and 5. In the meantime, however, we need to address the question as to whether there is adequate responsiveness in the system when it is in this structure.

§2.6.2 Reasons-Responsiveness

Does the restructuring of our agency cause the system to fail to track reasons, and does it cause us to lose regulative guidance? In order to still be a reasons-responsive system, and thereby exhibit the capacities required for autonomous agency, the answers to both of these questions must be no. As we saw in §2.3.3, mania alters the working of the salience and valence systems that allow our emotions to resolve the world into a normative picture, but not in a way that arrests these capacities. Manic moods still successfully resolve such a picture (as we saw in §2.2). This is a qualified answer, however. The system is capable of identifying reasons, however it takes reasons to be strict and decisive in cases where they are not. The system can represent our reasons to us, but not in a way that lets us engage in the usual deliberative weighing of reasons. This is a limitation, but not a fatal injury to our reasons-tracking capacities. Reasons-tracking does not require that we get relative weights right. Were this to be the case our emotions would very rarely qualify as reasons-tracking. Our emotional reactions often overrepresent the weight of a reason when the feeling is fresh, and underrepresent it when we have adequate distance. We might be angry that our bike was stolen last year, but the anger does not present the wrong to us as being as serious as it did when we discovered the theft and were furious. The correct weight of the offence is probably somewhere in-between. Despite this, these are all instances where our anger is successfully tracking reasons for others not to steal our bike.

Manic agents' autonomy, then, turns on the second question. Does this restructuring of agency, which narrows down our space for deliberation, mean that we are no longer committed to regulative guidance. I argue that it does not. Westlund has argued (2009) that an attitude of *self-answerability*, characterised as a disposition to give account for one's view of the normative landscape (in appropriate circumstances), is a crucial element of autonomy. Agents who do not have this disposition, who are untroubled by apparent reasons to doubt their view of the normative landscape, are engaging in some sort of passivity with respect to their normative outlook. To have this disposition, then, is to be non-

passive with respect to one's normative outlook, and to be sensitive to reasons to doubt that their outlook is reliable. In Tappolet's terms, agents who are self-answerable are agents who are keeping an eye out for the red light to go on, to indicate that their reasons-tracking system has stopped being reliable. I take it that the tendency manic agents have to wax lyrical about the merits of their projects and fixations exhibits self-answerability and indicates that regulative guidance persists. Manic agents are often disposed to speak at length, with enthusiasm, about how exciting the objects of their interest, with (or perhaps *at*) anyone who will listen. Indeed, the frustration manic people often feel speaking to neurotypical people who do not follow along with the sudden lateral moves and clang associations (or, as the manic person might see it, who are slow and sluggish, and obstinately mundane about everything) indicates an investment in being understood. This is the disposition to offer account that Westlund highlights. Manic agents whose tendency to wax lyrical, or rant furiously, about the merits or outrages of what concerns them are, it seems, exhibiting this attitude of answerability for their normative outlook, in a way that seems to suggest proper attention to whether their emotional perception of normative properties is reliable or not, and thus suggesting regulative control remains in place.

The manic agent keeps one eye out for the red light going on, but as we have established, they are still likely to be mistaken about the decisiveness of the reasons they detect. The optimism that is characteristic of many manic moods, and the "feeling of successful insight" (Lodge 2020) mean that regulative guidance is less likely to be *effective*. This commitment does not, in the end, guide us back to the right answer. Should we understand this to defeat, or at least endanger, regulative guidance in a way that would arrest our capacities for autonomy? Again, I contend not. Autonomy as an ideal is our excellent functioning qua agent. The capacity-cum-value that we exercise does not, indeed routinely cannot, meet that ideal. Getting our reasons right all the time is what we're aiming at, and anything that makes that harder (such as our regulative guidance being too light touch) is less than ideal, but this does not arrest the basic capacities, social recognition of which renders our agency autonomous. We can think of other problems for our agency that would interfere with good regulative guidance. Huckleberry Finn threw his trust behind a more reliable reasons-tracking system when his fondness for Jim indicated that something was wrong with his beliefs. He is uncommonly virtuous specifically because of his extensive moral miseducation. Any number of his contemporaries face the same systemic injustices that tells them that they are superior to other people. Endemic white supremacy,

propaganda for which was delivered in schools, churches, and government buildings, is the kind of force that interferes with the efficacy of agent's regulative guidance. Despite this, it would be strange to conclude that free people in slavery societies were not committed to regulative guidance simply because they were trained to be bad at it. It would be stranger yet to conclude that their basic agential capacities were arrested to the degree that social arrangements could not successfully confer autonomy on them. For any such individual, this is to invite disempowering explanatory answers to the question *why did they support slavery?* Seeking such answers casts Huck's contemporaries as vehicles through which a culture at a time acted, but not as agents themselves. This is to impose the explanatory injustice Mitova (2020) described. Rather, we should respect these agents by recognising their agential capacities, recognising their commitment to regulative guidance, (and recognise that part of the reason it was ineffective was societal), and ultimately, respect them by condemning their severe moral failings. Weak or misguided regulative guidance limits the degree to which we might excel as any kind of agent, but it does not arrest the basic capacities. Despite the restructuring of agency, there is still autonomy there to extend in various dimensions.

§2.6.3 Recognition

When social recognition is not forthcoming, this does not only leave our agency impaired by casting us as the vehicles by which social forces, or mental disorders, or the like act, it reduces the range of options in front of us. A great many acts, particularly those with social significance, cannot be conducted without our being recognised as the kind of person who can carry out these actions. Langton famously drew our attention to this fact when she argued that women's ability to issue sexual consent, or to refuse consent, is significantly impaired by a media landscape which constructs women as sexual resources available to men, and sex refusals as performative locutions rather than authoritative illocutions, (1993). The urgency that manic people respond to is typically not apparent to onlookers, yet despite this, when deliberative space is limited, the same swiftness of decision, and light-weighting of risks is called decisiveness. The manic agent might be decisive in circumstances that call for caution, but rather than being understood as reckless or even impulsive, manic agents are often subjected to disempowering explanatory answers to the *why* question.

Manic agents, insofar as they are subjected to disempowering descriptions, lose access to various actions. We have discussed the key example of this – Olga Blum was unable to recommend a book. The fact of her mania was taken by Binswanger to offer evidence of the workings of her manic grandiosity. I have similarly used it to motivate a view of how agency is restructured during mania. The step I have been careful to take is to retain a focus on Blum’s utterance as a piece of normative testimony. To ask ‘why did Blum say what she said?’ can be answered with reference to her mania, but also with reference to her view of reasons. She thought the book had such great value that the world should not be suffered to continue without it. From this testimony, we can learn about how Blum sees the normative landscape and how she assumes reasons are to be handled. Explanatory answers to the *why* question can inform us about how grandiosity works, but to understand how her *agency* has been reshaped by her mania we must first recognise her agency where it is. Blum’s view of the normative landscape is owed recognition if we are to inquire after her autonomy, as is the normative outlook of any agent whose autonomy we are curious about.

Insofar as this restructuring of autonomy prompts our interlocutors to take this disempowering, explanatory stance towards us, actions that rely on recognition of our normative perspectives are not options for us. To excel as an agent involves, among other things, exercising agency. When our options are reduced there are fewer ways to excel as an agent as fewer ways to exercise this agency are available to us. This limits the extent of our autonomy on at least one dimension of most theories, and might limit it on several.²⁹ There are, it seems, two different ways for mania to limit the degree to which our autonomy extends in various dimensions. It might limit our capacities for autonomous action, or it might prompt people to not recognise our agency as autonomous. Both of these limitations will apply differently in different dimensions - to use the topography metaphor, they take divots out of our autonomy in different places – but one kind of limitation is socially contingent rather than a given feature of the existential reality someone experiencing episodic mania must face.

²⁹ This limits our self-determination on Mackenzie’s theory (2014). On Killmister’s (2017), plausibly, this limits our self-realisation and our self-constitution.

§2.7 Conclusion

This chapter has argued that mania is frequently to be understood as involving a pervasive sense of urgency which restructures agency. This restructured agency changes not only the agent's normative outlook, making some considerations seem totemic and allowing others to fall into the background, but by making these totemic considerations seem strict and decisive, after the fashion of requirements. There are various ways in which this kind of agency will lead one to mishandle one's reasons, treating merely *pro tanto* reasons as if they were requirements, however there are also domains in which this kind of agency is valuable, either as a module that can be deployed to deal with genuinely urgent situations, or in which a manic agent might focus their efforts to ensure this structure can do the kind of work it is good at, and not disrupt the kinds of projects it is ill-suited to.

I have not gone into depth on the specific advantages of this manic structure of agency. That will be discussed at length in the next chapter. Rather, I have demonstrated the different sorts of impairment mania imposes on our autonomy. Some limit our capacities, however only to the extent that we are less successful agents, being less likely to respond appropriately to our reasons, but not by arresting those capacities nor making us stray so far from appropriate response that these capacities are unworthy of recognition. Manic agents will face a ceiling to how well they will excel with respect to various aspects of agency, and so a ceiling in how far their autonomy will extend in those dimensions. Chapter 4 will explore the role self-narrative plays in buttressing this kind of agency, and highlight the kind of self-narrative best suited to helping agents excel despite these limitations. Chapter 5 will examine a distinction in the scope at which self-determination might be predicated, which will help us distinguish between failures of autonomy to extend so far in this dimension, and two ways it might successfully extend in this dimension.

The other category of limits to autonomy in various dimensions are socially contingent. Rather than being compensated for by self-care practices that scaffold agency, these divots in our autonomy can be rectified by more empowering social relations. These kinds of relations are partly a matter of how our self-narratives play with our peers and how we appear to them on the other side of our relationships. The core of how this might be addressed, however, is by instituting practices of enquiring after mad

people's normative outlooks, and building greater understanding on both sides of these relationships. Chapter 6 will argue for a model of responsibility practice that creates these empowering social relations, specifically appropriate to manic agents.

Chapter 3 – A Moral Epistemology for the Aftermath of Mania

This is poetry as illumination, for it is through poetry that we give name to those ideas which are – until the poem – nameless and formless, about to be birthed, but already felt.

- Audrey Lorde

§3.1 Introduction

In the previous two chapters we examined the kinds of agency at work in depression and mania, which furnishes us both with a general model of moral psychology and an understanding of the changes in that moral psychology occasioned by affective shifts. We have pictures of different kinds of agency local to different parts of the complex warren that philosophy of agency continues to explore. However, one of the most distinctive features of mania and depression is that they are usually episodic phenomena. These affective shifts are disruptive, in part, because they move us onto terrain we are not familiar with, leaving us with kinds of agency we are not accustomed to. The particular agential challenge laid down by bipolar and related disorders (BoRD) is to develop adequate competency in the various kinds of agency we will find ourselves with over time. This chapter, as well as Chapters 4 and 5, can address particular features of that challenge.

This chapter focuses on the question of what to do with the experience of mania once it remits. Paul Lodge highlights a pressing dilemma that people who experience, or have experienced, mania face: how seriously should we take the manic sense of things? (2020) Should it be dispensed with, no more yielding than a dream? Or should it be interrogated for what it can tell us about ourselves and the world? What pressures come to bear on this choice, sharpening the horns of the dilemma? I will argue that agents who experience mania have sufficiently good reasons to take the manic sense of things seriously. It presents a valuable resource which is worth preserving in our curated inventories of agency, resources which serve moral-epistemic functions that, whilst not unique, are not readily carried out by neurotypical sorts of agency. The manic sense of things, therefore, earns its place by sensitising

agents to values they would likely otherwise overlook. This demonstrates, therefore, what excellence in manic agency looks like. The restructured kind of agency we explored in the previous chapter does its excellent work when it provides specialised modules of agency which are highly sensitive to otherwise obscure values.

This chapter will develop three arguments for taking the manic sense of things seriously. To take the manic sense of things seriously does not mean to defer to it out of hand, but to recognise it as normative testimony. We need not accept the testimony, but may not dismiss it out of hand, and will often benefit from synthesising it with other testimony. In §3.2 I will argue that not taking the manic sense of things seriously is a prohibitively unattractive response as it expresses and further cultivates a relationship towards one's own experiences that endangers one's agential capacities. §3.3 will outline the positive case in terms of two arguments: the argument from depression, and the edification argument, as well as specifying what must be the case for these arguments to be persuasive. The subsequent sections will provide arguments that satisfy the specified conditions. §3.4 will argue that there are values which the manic sense of things can be, and often is, responsive to. These are often *esoteric values*, that are difficult to grasp or articulate due to the limits of normal cognition. §3.5 will argue that mania tends to produce a fluency in thick evaluative descriptions, which allows the manic agent to grasp those esoteric values, due to the wider networks of meaning deployed, built on more relaxed standards of association. §3.6 will argue that something of the manic sense of things can be retained after mania remits, with reference to Nguyen's modular account of agency, introduced in the previous chapter (2020). §3.7 will reiterate the two arguments from §3.3, with all elements now in place, and demonstrate the value of the manic sense of things and the modes of agency that accompany it. §3.8 will conclude.

§3.2 The Aftermath Problem

Previously, we established that mania frequently occasions a restructuring of agency, according to which the world seems affected by pervasive urgency. The agent's attention is drawn to the normative

properties of certain things, and the normative objects of attention hold that attention with an iron grip. The reasons the agent is considering seem not only to outweigh other considerations, but to pre-empt them. When matters are urgent, deliberating about the weight of these reasons relative to the reason(s) that one's attention is locked on is neither necessary nor appropriate. This all feeds into a general sense of profundity; "at the heart of the experience is the feeling of successful insight." (Lodge 2020) The priority of the object of one's attention over everything else lends the experience an air of profundity, mysticism, or sanctity. Lodge describes this "as an ecstatic sense of the meaningfulness of existence, accompanied by the sense of an encounter with (though 'encounter' does not seem quite the right word) something 'sacred'. Yet, mania is an episodic phenomenon. When it remits, we are returned to a more familiar sense of what matters and how it matters, according to which we readily recognise reasons, value, and meaning in the world, but do not share the manic sense of things. This is the situation in which we face the *aftermath problem*.

§3.2.1 Lodge's Dilemma

The problem, as Lodge presents it, is as follows:

"How seriously should people who have felt this way take the echoes of what now seems like a distant land? Worse, the echoes of a land that is completely at odds with the one in which those around them seem to live? The dilemma becomes more severe given a social environment in which, if one were to take the manic sense of things seriously, one would run the risk of incomprehension, ridicule, and social alienation."

(Lodge 2020)

The person who has been manic has had powerful experiences of value, yet also has important contrary experiences of value. So far, this is not an uncommon position to be in. All agents have contradictory experiences of value, and synthesise them into a cultivated opinion on the matter. The difference mania presents is in degree rather than kind. The experiences are far starker in their content and in their difference, they present far greater sorts of value as knowable, and they pull many more nodes into a complex network of meaning. This difference in degree is sufficient to make manic perspectives resistant to interpretation by non-manic people. Lodge's dilemma can be expressed in a number of ways. Does this experience offer moral knowledge? (Does it offer anything that is even a candidate for moral knowledge?) Should the recently manic agent offer the experience to others as moral testimony,

or accept their memories of it as moral testimony? I contend that the manic sense of things does offer moral knowledge, however it is important to understand the available options before I proceed with this argument.

§3.2.2 Sharpening the Horns of the Dilemma

People who experience episodic mania face an invidious choice, between alienation from their own experience, that might threaten to eliminate much of their agency, or enduring the social identity ‘mad person’, which strips one of epistemic agency and the standing to contribute practically. They remember a normative outlook according to which some values were totemic but could also pull in a near unlimited number of related objects or ideas, due to the eternal network that manic ideation tends towards. The first horn of the dilemma threatens us with alienation from our own experiences of value, which is more than usually costly in the context of mania due to the intensity of the experience. On the other hand, Lodge highlights the social costs of being seen to take the manic sense of things seriously. Agents who do so invite unintelligibility, ridicule, and social alienation, but particularly, they risk occasioning a disempowering pathologizing stance. We have already discussed Mitova’s concept of explanatory injustice (2020) which occurs when others explain what we do or say with reference to disempowering explanatory terms (such as our culture, mental disorders, or ‘feminine emotionality’), rather than with reference to our agential capacities (our view of which reasons matter and what response they call for). When our behaviour or speech is pathologized we are disempowered as agents as, like the victims of explanatory injustice, we are cast as the vehicles by which a mental disorder acts on the world.

In 3.2.3 I will argue that the first horn of the dilemma is prohibitively unattractive, which will serve as the first of three arguments for taking the manic sense of things seriously, but before that, I will briefly consider whether I can blunt the second horn for the sake of a softer landing.

At a fundamental level, the dilemma Lodge presents is the same dilemma that Srinivasan describes as an *affective injustice*, (2018). These are the injustices created when victims of oppression must choose either to suppress their legitimate anger for the sake of making progress in overturning their oppression

or expressing that anger and variously deepening their oppression or making it even harder to end than it already is, (127). The anger of the oppressed often induces a disempowering explanatory stance in others. The prejudicial social identity of ‘the angry black woman’ is used to explain why someone is angry, rather than her sensitivity to reasons of justice, the latter of which risk presenting her anger as moral testimony in need of recognition.³⁰ In order to do anything about oppression, the oppressed must frequently tell their oppressors what they are doing. If anger prompts disempowering explanatory stances towards the speaker, their testimony is lost in the exchange. Thus, in order to do anything about their oppression, speakers must be careful not to let their anger trigger these defensive responses. They must alienate themselves from their experience of the intense disvalue of the way they are treated. At its heart, this is the dilemma that post-manic agents face. ‘Mad person’ is a stigmatised and marginalised social identity that is particularly susceptible to the kinds of epistemic subordination that protect systems of privilege in other domains. Whilst the person who is in the aftermath of mania might not, at that moment, have serious sorts of oppression to vie against, their readmission into any epistemically dominant (non-mad) identities is frequently contingent on their alienating themselves from their potent experiences of value.

To blunt this horn, it might be objected that the manic sense of things is too unlike righteous indignation for this analogy to hold. *Ex hypothesi*, the anger that Srinivasan considers is apt anger, which fits its object. The manic sense of things, on the other hand, might be thought to be unfitting, casting values in all the wrong proportions and leaving one with a quite mistaken view of the normative landscape. Firstly, I contend that the manic sense of things often does proffer moral knowledge, which I will argue for in §3.4. But secondly, both cases share a sort of alienation as a harm. This is the fundamental harm in Lodge’s version of the dilemma but is a secondary harm in Srinivasan’s formulation. Being required not to properly feel apt anger is bad for us because anger *appreciates* injustice in the correct way and we are better off insofar as we appreciate things that are to be appreciated, (132).³¹ Agents who suppress their apt anger might be effective at testifying to their oppressors in this or that instance, but as

³⁰ Compare this to Alison Bailey’s account of tone-management as a method of silencing morally important testimony, (2018, 97). The legitimate anger of black women is silenced when hearers focus on the tone of delivery as a violation of social etiquette, rather than as a feature of the particular kind of moral testimony that is delivered.

³¹ This is a more substantial claim than the sentimentalist thesis, as it specifies not only that emotions *can* represent to us the reasons for which we eventually act, but that we *should* emote in the relevant way in the face of certain reasons or values. To coolly judge that an injustice is heinous is to fail to fully grasp that it is, (compare Wolf, 2015, 112).

this is required more and more over time, their relationship with their own experiences attenuates. The dilemmas share alienation as a key harm, differing only in whether it is an immediate or non-immediate harm. Irrespective of which, alienation makes its horn of the dilemma unattractive.

It seems, then, that I cannot blunt this horn of Lodge's dilemma. What this comparison does show, however, is that the social sanction horn is different in kind from the alienation horn. As I will argue under the next heading, alienation is bad for our agency due to the way it weakens our relationship with ourselves. The social sanction horn, however, is painful only insofar as it is made painful by unjust social arrangements. Taking the manic sense of things seriously does not disempower us by interfering with our agential capacities. As demonstrated in the previous chapter, there are distinctively structured agential capacities there to recognise. Rather, it is made disempowering by other people whose familiarity with pathological terms, but not the manic sense of things itself, prompt them to take disempowering stances towards agents who perform the social identity of 'mad person' by being seen to take mania seriously. This will not make the latter horn less painful when it hits, but the choice is between throwing oneself on the former horn, or letting others strike one with the latter. Neither is attractive but given the damage to one's agency that this kind of alienation can produce, we at least ought not do this to ourselves. We do not solve socially created problems of this sort by acquiescing to them, we solve them by confronting them with robust agency.

§3.2.3 Alienation and the Tentative Arrest of Agency

This alienation damages agency insofar as it breaks down the agent's sense that their own normative outlook is reasonable. Agents rely on a sense that they are, overall, competent to recognise worthwhile reasons in order to carry out an effective agential project and to avoid autonomy-detrimental passivity or deference.³² To lack this sense of one's own competence endangers all our moral knowledge. To think that 'I believe that ϕ is valuable, but that is because I have an unreliable sense of these things'

³² Govier (1993) calls this self-trust, without which we would struggle to function as agents. She remarks that: "To lack general confidence in one's own ability to observe and interpret events, to remember and recount, to deliberate and act generally, is a handicap so serious as to threaten one's status as an individual moral agent." (Govier 1993, 108, compare McLeod & Sherwin 2000). We shall explore this theme more in the next chapter.

provides what McGrath calls an undercutting defeater for moral knowledge, (2019).³³ We can lose moral knowledge, she argues, when we encounter evidence that we hold beliefs for reasons unrelated to the truth of those beliefs. In empirical cases, she offers the example of believing that a coin landed heads last time it was flipped, which is knowledge (it is true, and you remember it happening) until it is undercut by evidence that this memory was implanted in you by a hypnotist who has no idea which way the coin went, (172). On discovering this evidence, you lose knowledge about which way the coin landed because your justification for the belief is undercut by new evidence. Moral knowledge, she argues, is also susceptible to this kind of undercutting by ‘debunking explanations’. We have, she tells us:

no a priori guarantee that evidence will not emerge about why we hold our moral views that would undermine the rationality of continuing to endorse those views. Our moral beliefs, like our ordinary empirical beliefs, are susceptible to being debunked if the right kind of evidence emerges.

(McGrath 2019, 190, original italics).

Evidence might take the form of Marxian analysis of the interests of emerging loci of power, psychoanalytic analyses of human drives that need to be managed in equilibrium between people, etc. There are always a range of factors that can play an explanatory role as to why we hold the beliefs we do, whether about morality or about something else. One such explanation that would debunk moral knowledge is that ‘I believe *p* because my cultivated faculties bias me towards such beliefs’. This is probably true for us all to some degree. Our cultures and social positions do bias us in various ways. Our susceptibility to advertising and propaganda will prime us to accept some bad premises in our moral thinking. A man might believe that men and women are largely paid what their labour deserves and be uncomfortable with the idea that he is complicit in unjust systems, or that his success is not wholly his own. Nonetheless, he would do well to reflect on the role these feelings play in stabilising his belief about wage labour. A critical awareness of these debunking explanations can be, locally, extremely valuable in securing greater moral knowledge.³⁴

³³ We considered rebutting defeaters in §1.5.3.

³⁴ I assume, however do not have the space to argue, that debunking explanations that have all moral knowledge in their scope (evolutionary accounts, Marx’s account of ideological superstructure serving to stabilise an economic order, etc) do not eliminate the possibility of moral knowledge.

Yet, despite the importance of attention to debunking explanations of our moral beliefs as a local corrective, it should not be our general stance towards our own agency. To decide not to take the manic sense of things seriously is to adopt such a stance, taking one's mood to explain one's moral beliefs, thereby undercutting any (candidate) moral knowledge one enjoyed during that mood. The deep problem with this attitude is that we are always in one mood or another, and it is not at all obvious that mundane moods cannot offer undercutting explanations if manic moods can. Recall, from the previous chapter, Radden's analysis of moods as disclosing the world under a particular normative description, (2013). Manic moods, variously, disclose the world as brimming with possibility, or agitating and uncomfortable, or full of outrages and personal offences. If the manic mood is an undercutting explanation of the manic person's moral beliefs (candidate moral knowledge), then mundane moods, which present the world under more familiar normative descriptions, are surely also undercutting explanations. They hold these normative beliefs because their (mundane) mood disposes them to, not because of the moral-epistemic status of the beliefs. One might reject mundane moods' status as undercutting explanations by understanding moods as emerging from the normative features one is sensitive to, as outlined in §2.2.1, and therefore not independent of the agent's view of reasons. Due to this lack of independence, mundane moods cannot be undercutting explanations of one's moral beliefs, because to talk about moods is simply a zoomed-out, lower resolution, way of talking about the agent's many moral beliefs or perceptions at once. However, this is as true for manic or depressed moods as it is for mundane ones. The problem with seeking out undercutting explanations is that we are likely to find them all too often. Seeking undercutting explanations in certain contexts is a valuable corrective, but a general policy of systematic doubt towards one's experiences of value will undercut all of one's moral knowledge.

Such a policy of doubt costs more than our moral knowledge. The approach here is well captured by the attitude of *tentativeness* that Talia Weiner explored as a self-care practice shared by a number of attendees at a bipolar peer support group (2011). Weiner tells us that:

As invoked by the group on several occasions, tentativeness referred to... a cultivated stance of uncertainty and suspicion toward one's own thoughts and emotions at any given moment. Tentativeness thus described a distinct disposition in relation to risk—one that required an acknowledgment of the management of future uncertainty by an uncertain present self. Discursively enacting tentativeness by expressing awareness of one's own uncontrollability and skepticism toward the legitimacy of one's own

emotional experiences indexed members' expertise at self-management by demonstrating a kind of vigilance.

(Weiner 2011, 472)

This attitude is, at its heart, the attitude that we must take towards at least our manic experiences if we throw ourselves on the first horn of Lodge's dilemma. It risks costing us moral knowledge, but Weiner highlights a further threat - this tentativeness can wholly arrest one's agency. An especially articulate participant in Weiner's research, James, gives us insights into how this works. His cultivated scepticism towards his mental states left him unable to determine whether an idea he had intended to pursue (to write an article) was properly his. He lost interest in the idea as he adjusted his medication, and wondered whether it was an instance of "losing a thread of mania? Or is that the drugs completely, you know, just sort of squashing me flat?" (470) He is seeking to distinguish mania (reified as some alien force) from his own agency, but relies on external comparisons. He remarks that he doesn't know whether other people in his position would be comfortable with "basically doing nothing? Which is mostly what I do with my day, is nothing." (471) James' might have been an extreme expression of tentativeness, but it is an expression that wholly arrests his agency.³⁵ Cultivated suspicion towards one's normative outlook, on the grounds that it *is one's own* normative outlook, presupposes that the agent is not competent to recognise reasons, and seeks out undercutting explanations everywhere. Mania is susceptible to undercutting explanations, as James puts it: "bipolar alone is not sufficient to explain every behaviour that you do—but neither is it not the explanation for everything that you do." (470). But once undercut in this way the agent loses self-trust. Second guessing all of one's own motivations is a recipe for much the same inertia as is futility thinking.

It might be objected that, whilst we should not take an alienated stance towards our normative outlooks generally, as undercutters will be too readily found, that mania is one case in which we should look for undercutting explanations because we do need local correctives. Whilst manic moods are, qua moods, very like mundane moods, they are distinctive in some relevant way. This relevant way might be that they are pathological moods, as opposed to non-pathological moods, or it might be that manic moods

³⁵ Tentativeness can play supportive stabilising roles in one's agency when deployed carefully. We return to this example in §5.3.1.

(irrespective of pathology) happen to be morally incompetent moods - moods that simply fail to adequately portray the normative landscape. I will address these objections in turn.

To seek out debunking explanations as a local corrective on the grounds that manic moods are pathological is subtly circular. The description 'pathological' comes to be appropriate or inappropriate only when we take an explanatory stance towards the phenomenon in question. Pathological explanations are a specialised sort of explanatory stance that ask questions to which the agent's reasons are not the answer. If we take agentic stances, we simply cannot trade in terms of pathology as we are asking different questions. To seek an undercutting explanation on the grounds that mania is pathological is to start with an explanatory stance, and then seek an explanation rather than an agentic account ('what brought about this agent's behaviour or speech?', rather than 'for what reason did they do/say it?') At heart, the demand to find undercutting explanations on the ground of mania's pathological status is a demand to find explanations because there are explanations to be found. In empirical enquiries this is precisely the thing to do, however *this* enquiry is concerned with recognising, on its own terms, the agency that is present in atypical mood states. Recognising this agency requires an agentic, rather than explanatory stance. Indeed, agency is only visible when we take this kind of stance. Whilst a great deal of interesting and useful knowledge can be gained about our brains and bodies, our experiences, and our psychology by taking the pathologizing explanatory stance, the question of how an agent living in the aftermath of mania should think about those experiences is governed by the need to avoid alienation and tentative arrest of agency.³⁶

The second ground for this objection is that mania, due to the restructuring outlined in the previous chapter, is likely to mislead us on the balance of our practical reasons in a significant number of cases. Particular normative properties of things hold our attention with an iron grip, and seem to have priority over other considerations, to the degree that the agent need not or ought not deliberate further. In a great many cases this will lead agents astray. I respond on two fronts.

³⁶ In a similar vein, Radoilska (2022, 169, see also 2013) suggests that "necessarily less-than-successful-agency" is a concept that can occupy at least some of the logical space given over to pathologies of agency understood in clinical terms. Taking an agentic stance, in this way, casts feature of mental disorders as constraints to agency that can be worked within, or overcome, rather than intrusions of non-agency.

The first is that this provides agents good reasons to keep undercutting explanations in the periphery of their vision whilst manic (cultivating strong habits of self-management that highlight the need to check one's sums over) but provides such reasons only during mania. The question addressed in this chapter is how agents should relate to their manic experiences in the aftermath of mania. Roughly, the distinction is between the governing norms – right action in the midst of mania (when the manic sense of things is highly likely to determine what one does) and right belief in the aftermath (when that sense is one prominent candidate among many memories we have of other normative outlooks). The second front is that we can take the manic sense of things seriously without acceding to the kinds of mistakes it tends to produce. Taking the manic sense of things seriously requires only that we recognise it as moral testimony, and a possible source of moral knowledge. Whether it is good testimony, or a source of real knowledge, can only be determined after we extend this recognition. I contend that the manic sense of things will tend to produce a specialised kind of good testimony, and so we should take it seriously both to avoid an unhelpful alienation, but also so that we can gain esoteric sorts of moral knowledge that we might otherwise not get the chance to acquire.

The first horn of Lodge's dilemma is prohibitively unattractive. What, however, can be said in favour of accepting the second horn? The next section offers two related arguments that make up the positive case.

§3.3 The Positive Case

The argument from alienation and tentativeness applies strong normative pressure against accepting the first horn of the dilemma, however is not a knock-down argument in its own right. There could be stronger reasons against taking the second horn and, if the second horn offered no benefits, it could be at least as unattractive. It is not practical to consider everything that might be a reason not to accept the second horn and demonstrate them to be insufficient, but what is achievable is to demonstrate that there is particular value to be found in taking the manic sense of things seriously. In this section I will outline two related arguments, and make clear what needs to be the case for these arguments to be

effective. The first argument holds that taking the manic sense of things seriously can make depressions significantly more survivable, and the second argument holds that the manic sense of things is morally edifying in its own way, and so worth taking advantage of.

3.3.1 The Argument from Depression

People who experience mania are very likely to also experience depressions. This is both because depressions are ubiquitous experiences in most human communities, but also because people who experience mania (at least as far as the clinical concept goes) tend to experience it as part of a bipolar or related mood disorder (ICD-11 6A6), distinctive of which is alternating episodes of mania, hypomania, and depression. It is not unheard of for people who experience these conditions to consider their mania either a kind of compensation for their depressions, or as something that secures meaningfulness in their life in opposition to the nihilism that their depressions suggest. This argument suggests that the manic sense of things should be taken seriously as it is a particularly effective way to rendering depressions less harmful or more survivable. The thought is that the particular depressive syndrome, where one's experiences of positive value dry up and one's judgements are eventually trained to endorse a nihilistic outlook, could be slowed or arrested if one habitually took the manic sense of things seriously.

The animating thought behind this argument is that depressions are bad for us, at least in part because they blind us to the value of certain things, or because our lives fail to seem meaningful to us. The manic sense of things, on the other hand, gives us an intense sense of value and meaning. By taking the manic sense of things seriously, therefore, we have a bulwark against depressions that will make them less painful. This is not to say that taking the manic sense of things seriously will prevent a significant number of depressions from coming on, but rather that it will give agents resources with which to better endure depressions. The process described in Chapter 1, where the agent's moral sentiments become less and less sensitive to goods (yet might remain highly sensitive to bads) had the consequences of training their judgement into a similarly nihilistic state. What I suggest is that a cultivated habit of thought, whereby agents take the manic sense of things seriously as moral testimony, might slow this rate of decay. At the margins, this will avoid a few depressive episodes from emerging, but is valuable more in softening rather than averting depressions.

This is a very old argument. It, or something very like it, is implicit in Plato's discussion of the divine madness that Dionysus bestows upon people who are afflicted with bloodguilt (a heritable spiritual pollution that makes one sick and non-Divinely mad). Dionysus' mystic inspiration is one of four sorts of divine madness he discusses in the *Phaedrus*, along with prophesy, poetry, and love. Each of these serve our moral edification in some way, giving us insight into virtue or the value of things, and prompting us to reflect in ways that will unlock knowledge of the Forms for us. Of Dionysus' gift, he tells us:

[M]adness enters certain famous families which have been afflicted by horrendous illnesses and suffering as a result of guilt incurred some time in the distant past, and with prophetic insight finds the necessary means of relief. It resorts to prayer and worship of the gods, and when, as a result of this, it comes up with purificatory rituals, it makes the madman better, not just temporarily, but for the future too. And so it finds a way to release people who are mad and possessed in the right fashion from the evils that afflict them.

(Plato 2002, 244d-e)

I understand mystic insight to involve something like an intense experience of the spiritual significance of things, in many ways analogous to the intense experience of value and significance involved in mania. Intense experiences of these things are, Plato suggests, effective ways of overcoming one's afflictions.³⁷ Taking depression to be a reasonable candidate for the kinds of afflictions in question, and the analogy between mystic inspiration and the particular features of mania I have focused on here, I take it that this argument is a developed form of the one Plato offered.

What must hold for this argument to be persuasive? Firstly, the manic sense of things must somehow be preserved. It is not obvious that this will be the case. Whilst mania involves a powerful sense of meaning or value, it might be a sense that is also susceptible to the erosion that depression brings about. An argument must be offered to support my assumption that it is not. It shall be developed in

³⁷ I have been able to find no philosophical work examining Plato's attitude towards Dionysus in this context, which is unsurprising as it is one fleeting example of among several, offered to situate *love* as a species of divine madness, which is his focus in this part of the *Phaedrus*. More work has been carried out in the Classics literature, where the consensus seems to be that the rites Plato has in mind serve much the same role that placebo or psychosomatic treatments would serve in our medicine, (Waterfield 2002, 86, citing Parker 1996, 288). Scullion, (1998) argues, with reference to this passage (244d-e), that it is chiefly mental distress rather than physical for which Dionysus would be invoked.

§3.6. Secondly, this argument invites the following challenge: perhaps taking the manic sense of things seriously is useful for making depressions more survivable, but if we are simply self-deluding then we have made a trade we perhaps ought not to make. Depression can be more survivable, but if this comes at the expense of truth or moral knowledge, it is not obviously a worthwhile trade. If we are self-deluding, it might be thought we are living a meaningless life. Indeed, this is what had been going wrong for the people Wolf describes who change their mind about what seems meaningful and who are right to do so (2015, 132). If we adopt her slogan: “meaning arises in a person’s life when subjective attraction meets objective attractiveness” (ibid) then trades at the expense of moral knowledge as to what is objectively attractive are unworthy trades. We might suffer less depression, but no longer enjoy meaning in life. In response to this, I argue that the manic sense of things contains something veridical, (§3.3.2, below). It gives access to moral knowledge that is otherwise difficult (although perhaps not impossible) to acquire. Thus, taking the manic sense of things seriously is not a trade against meaningfulness, it is simply (like every other pursuit of meaning) susceptible to being mistaken.

3.3.2 The Argument from Moral Edification

The second argument holds that there is something veridical in the manic sense of things, and that something would be otherwise difficult to attain. The argument goes that mania sensitises us to a family of reasons, values, sorts of meaning, or normative properties which we struggle to grasp or articulate in non-manic states. Thus, mania offers advantage in the pursuit of moral knowledge. The only way to secure the manic sense as a source of this moral knowledge is to take the manic sense of things seriously. Thus, we should take the manic sense of things seriously. This argument is necessary to foreclose the meaningfulness objection to the previous argument but if it is persuasive, then in its own right it applies normative pressure in favour of taking the manic sense of things seriously.

The animating thought behind this argument is that standard, neurotypical, or non-mad sorts of agency are limited in a number of important ways, and that these limitations do not overlap wholly with the ways in which manic or other sorts of mad agency are limited. The lack of overlap means that different kinds of agency can be called on to cover the gaps left by others. The sets of agencies that we cultivate

amount to what Nguyen called a private *inventory of agencies*, (2019, 2020).³⁸ The relevance here is that the manic sense of things allows us to grasp certain values or sorts of meaning that non-manic kinds of agency do not allow, and by taking the manic sense seriously, we retain that sensitivity as part of our curated inventory of agencies. Such an inventory of agencies, more sensitive to these *esoteric values*, is one that grants an agent privileged access to particular kinds of moral knowledge. The manic sense of things, therefore, enables a narrow but specialised sort of moral edification, and is valuable insofar as it does so. It is valuable insofar as it offers esoteric moral knowledge, and valuable insofar as it allows us to track reasons which we would otherwise likely have been insensitive to.³⁹

What must be the case for this argument to hold? Firstly, and as above, it must be the case that the manic sense of things can be preserved or held onto after manic episodes. Attention to Nguyen's theory of inventories of agencies will demonstrate this in §3.6. Secondly, there must be something veridical in the manic sense of things that is otherwise hard to come by. We must be in a better position to gain some kinds of moral knowledge, to appreciate particular kinds of value, or participate in particular kinds of meaning, when engaged in manic agency than typical non-manic agencies. There must be something veridical and rarefied in the manic sense of things. §3.4 will argue that this is the case.

§3.3.3 These arguments are related

The argument from depression might be satisfactory on its own if we are unconcerned about meaningfulness in life or do not believe that false moral beliefs (of the sort mania might produce) do enough damage to someone's wellbeing to outweigh the advantages of being more resilient to depressions. If, however, we believe that the kinds of moral knowledge that mania might mislead us on is crucially important, or we take meaningfulness seriously along Wolf's lines, then we will require the argument from moral edification to be persuasive to foreclose that objection. The edification argument

³⁸ We discussed Nguyen briefly in §2.4 and go into significantly more depth in §3.6.

³⁹ This isn't a form of situated knowledge, as usually understood. The manic person is likely to occupy a social identity during their mania, either a subordinating identity like 'mad person', or an empowering identity like 'visionary' (but note that the later identity can be a mixed blessing, see Jamieson, (1994, 4-5) concerning the poet Robert Lowell). These social identities are, if the situated knowledge theory is correct, likely to offer privileged access to certain kinds of knowledge. However, it is not in virtue of these social identities that the manic agent has privileged access to the esoteric moral knowledge we discuss in §3.4, and so this is not situated knowledge in the usual sense.

is independently important as it demonstrates value in the manic sense of things that is not conditional on something like depression. Rather, insofar as the manic sense of things is valuable in our moral epistemology, it is useful to all agents. Those who experience mania have good reason to take the manic sense seriously, and others have good reason to take on board moral testimony that arises from the manic sense of things.

These arguments turn chiefly on the same premises. For either to be persuasive, we must have some way of storing what we learn from the manic sense for later use, when the manic experience of reasons has passed. This is necessary so that we can access the manic sense of things in the face of oncoming depressions, and so that we can appreciate the value of things we were sensitive to during mania after it passes. For the second argument to be effective, or the first argument to evade the meaninglessness objection, two further things must be the case. One is that the manic sense of things sensitises us to value, there are at least some kinds of value or meaning that the manic sense of things portrays accurately or appropriately. The other is that we are better able to grasp these values as a result of mania than they would otherwise be. The remainder of this chapter will argue that these desiderata are met.

§3.4 Mania Sensitises us to Value

In the previous chapter we established that the manic restructuring of agency causes agents to take certain reasons as if they were requirements. The reasons one is focused on seem to enjoy priority, making it not only unnecessary, but perhaps inappropriate to deliberate further. Matters seem urgent, even when they are not urgent. This leads manic agents to fail to respond appropriately to the practical reasons that speak to them. It is easy, therefore, to imagine mania as damaging to the excellent functioning of agency. Whilst we might fixate on good reasons, the nature of the fixation causes us to fail to appreciate that reason's status amongst other reasons, and so mania will systemically mislead us. Despite this, I argue that the kind of appreciation involved in this structure of agency is useful for acquainting us deeply with particular values, especially those values we otherwise struggle to grasp. In

this section I will outline what these values are and why we struggle to appreciate them, and then argue that distinctive features of manic cognition are useful ways of helping us better grasp those values.

§3.4.1 We systemically under-weigh some reasons

We are limited and imperfect agents, and so whilst we will aspire to excellent recognition of and response to reasons, we will never perfect these capacities. This, being uncontroversial, entails a minimal version of the proposition in this subheading: we systemically under-weigh some (families of) reasons. This is systemic insofar as our reasons-responsive systems admit of these limits. I want to motivate a stronger version of the claim, according to which something about particular (families of) reasons, or about the environment in which we attempt to detect these reasons, makes them harder for us to grasp, articulate, or properly appreciate. I call these *obscure* or *esoteric reasons*, grounded in *obscure* or *esoteric values*. I consider three families of esoteric reasons. There are the reasons we are trained to ignore, the reasons we do not have the conceptual vocabulary to grasp, and the reasons that are simply hard to pin down. This last family is the chief example I will have in mind, but my argument applies to all of them.

The first family are those reasons we are trained to ignore. On Nguyen's analysis, we slip into layers of agency to play games, which equip us to recognise certain facts as reasons calling out for particular kinds of response (2019). Much the same could be said of one's profession. Different jobs trade in different reasons, and so sensitise us in different ways. Importantly, we spend a great deal longer submerged in these layers of agency than we do in games. Consequently, they are likely to contribute a great deal more to cultivating our overall agency. If, however, some consideration does not count as a reason as far as our profession goes, this cultivation can make us increasingly insensitive to that consideration. This might be particularly more likely if the consideration called into question the legitimacy of what we were engaged in. We could readily understand if a prosecution barrister was dismissive of reasons to pursue moral repair and avoid carceral solutions to social problems. It would not be virtuous, but we could understand why they made a habit of ignoring considerations that drew their attention to the failures of carceralism. This barrister comes to systemically under-weigh reasons grounded in the value of forgiveness and repair, because they prefer not to imagine themselves as complicit in unjust or ineffective systems. Berman's recent history of the 'economic style of reasoning', and its spread

through American governmental departments, offers a model of this kind, (2022). This style prioritises efficiency and has a normative vocabulary in which all values are commensurable. The problem, of course, is that many values are not commensurable with each other and are not expressible in the terms of this style. The value of things like clean air and stable biospheres were systemically under-weighted because economic reasoning has limited means to measure (proxies for) these values and express them in its own terms. We can come to systemically under-weight some reasons if we submerge ourselves too deeply for too long in layers of agency that block them out. These values are obscured by the kinds of agency we cultivate.

The second way in which we come to systematically under-weight reasons is because we lack the (thick evaluative) concepts required to capture these values. This might be because there are particular evaluative concepts necessary to pick up the value in something which a community does not have access to, or because the concept is not understood to have this evaluative content. 'Sexual harassment', as a concept, belonged to the former category. Prior to the emergence of this concept, women who experienced sexual harassment did not have hermeneutical resources adequate to make sense of the experience, or why it was so distressing, (Fricker 2007, 150-1). A powerful but vague sense that one ought not be treated this way was insufficient for affected women to properly appreciate *what* was going wrong. In many cases, women would accept the explanation that they were excessively sensitive, and did not have a legitimate complaint, (compare, Jagger 1989, 166). Cases such as these are where we under-weight reasons due to some hermeneutical gap, a lacuna in the interpretive resources we inherit from our community. This can be closely related to the obscuring of values. The set of hermeneutic resources that the economic style of reasoning trades in, according to Berman's history, became the received set of hermeneutic resources, and so all those who entered the community that used this style (government technocrats, bureaucrats, legislators, jurists, etc) inherited the hermeneutic gaps of the style. The training we receive for a particular job or to face a general labour market might train us to under-weight, but if this is generalised across a community, it can begin to impose hermeneutical gaps that effect even those who have not received the relevant training. These values are obscure because they sit in our hermeneutical gaps.

The final family of esoteric reasons I consider are difficult to grasp due to something about the nature of the values that ground them. I will call these *ecological values*, (ecological in the sense of things living and thriving in community with each other, rather than the sense of being about biospheres).⁴⁰ Ecological values are esoteric, I suggest, due to their being grounded in something a little out of reach of normal human experience. There is overlap with the reasons obscured by hermeneutical gaps, but whilst a richer conceptual vocabulary is likely to make it easier to grasp ecological values, there is some limitation to how much easier it can get. It is the value, rather than the conceptual vocabulary expressing the value, that makes it difficult.⁴¹ The example that comes most readily to my mind is that of singing and whooping as one dances around a bonfire. There is (I insist) something fundamentally good for the soul in this practice that is not reducible to the pleasure it creates, the sense of community it might foster between humans, or the continuation of valuable traditions that it might carry on, etc. Being so esoteric, it is not something I could effectively argue for. Nonetheless, *if* I am right, then we are better off for knowing this. We have moral knowledge that allows us to appreciate this value to at least some minimal degree (not as well as we might appreciate non-ecological values, but more than not at all). Applying Wolf's formulation, we are able to find meaning in this practice and not be mistaken. Our subjective attraction to the practice meets the objective, though opaque, value in the practice. More familiar examples might make the case more persuasively. Ample literature in environmental ethics addresses the question of how we might be correctly oriented towards nature, avoiding the trap of anthropocentrism, which only recognises human value in the world, or at the extreme, takes the natural world to be of value only *to* humans, (eg, Hursthouse 2007, Altshuler 2014, Bannon 2017). The difficulty this literature confronts is that our being properly oriented towards nature (Hursthouse's formulation, 2007, 164) combines what is proper for us, human beings with particular limits on our perspectives, and the value that exists in the natural world that is outwith our ken. We cannot fail to be properly oriented towards nature because we do not properly appreciate things that are simply too inhuman for us to

⁴⁰ I am fond of novelist Susana Clarke's rendition: "the idea that the world is important because it is *not* human, it's something we might be part of a community with, rather than just a resource." (Jordan, *The Guardian*, 2020). We might variously call these values peripheral, inhuman, spiritual or mystic, or even pantheistic values. I opt for 'ecological' to emphasise the communion found in properly recognising and appreciating these values.

⁴¹ Some scholars will doubt that such values exist, holding that hermeneutic gaps are adequate to explain the difficulty we face. If they are correct, then under-weighting due to hermeneutic gaps will be the lead example my argument applies to, and it can be substituted easily for ecological values. I will, however, briefly say a word in defence of this thought. I start with the premises that the world is normatively rich, and that human experience can at best offer only a relatively narrow window on the world and its value. The first of these is controversial but I do not have space to defend here, the second should be uncontroversial. It would seem to follow that there is a great deal of value in the world that is on the margins of, or outwith, human ken.

grasp, but we feel the demand to try because there are values that are just on the periphery of our human experience. We better appreciate what there is to be appreciated, that is to say, we enjoy more moral knowledge if we can grasp these ecological values, even loosely. The virtue(s) that Hursthouse seeks are those that would see us related to the natural world, on its own inhuman terms, with the right feeling and thought, in the right proportion, on the right occasion, where the rightness standard accommodates the limitations of human ken. If we can force our ken a little wider, however, there are more values to appreciate and we are better off if we can appreciate them. We are better off because there are more ways our lives can be meaningful, and because we enjoy greater moral knowledge. Part of what it means for us to be agents is that we *appreciate* value where it exists, and our agency functions more excellently the more our agential capacities recognise value, even if it is esoteric, obscured, or peripheral.

These esoteric values are likely to be controversial, at least in the role I have assigned them here. I do not have space to offer a fuller defence of the notion. If one rejects ecological values, one is likely to hold that hermeneutic gaps are adequate to account for the cases I refer to: between them, Orphism and Sufism furnish us with concepts that are, *mutatis mutandis*, adequate to grasp the value of dancing and whooping around a bonfire. As both the philosophical work in environmental ethics, and the vast lived experiment on the subject both continue, we develop better conceptual vocabulary with which to appreciate the value of the natural world. Yet many hermeneutic gaps persist. If one is persuaded by this account, then filling hermeneutical gaps becomes the particular task to which the manic sense of things can contribute. As I am persuaded that there are ecological values, I take it that the manic sense of things can be useful for both. The following section provides the argument to this effect.

§3.4.2 Manic Insight

Mania gives us insight into esoteric values by virtue of two features. The first is the attentional effects of the restructuring of agency discussed in the previous chapter, and the second is the cumulative effect of certain features of manic cognition. I will address the former here and the latter in §3.5.

In Chapter 2 we established that mania occasions a particular restructuring of agency. The reasons one's attention is fixed on seem to enjoy a priority that makes considering the other reasons either unnecessary or inappropriate, as if matters were urgent and there was no time or space for deliberation. This fixation, even when it runs into obsession, is what deeply acquaints us with the value in question. The weakness of this attention is that it ignores other reasons that are relevant in the setting, however having gone from ignoring the value or reasons one is fixated on to ignoring other ones, we should recognise this as an overcorrection, rather than a new mistake. It is an *overcorrection* because the manic agent is unlikely to strike the right balance between this new value and familiar values well, but it is an *overcorrection* because the agent has come to recognise value that they previously did not. This can create problems for in-the-moment reasons-responsiveness, as we discussed in the previous chapter, but if we zoom out and look at this as a step in the continuing process of becoming better attuned to the value of things, then this is a positive step. Focused attention is what acquaints us with the reasons we are focused on. Like Nguyen's games, manic obsession "digs narrowly, but digs deeply" (2020, 88). Unless there is reason to believe that the manic sense of things systemically presents disvalue as value, this is a perfectly acceptable instance of familiarising oneself with values or reasons that are rather novel to one.

The process can better be appreciated if we look at our examples of obscured reasons or esoteric values. When it is our training that dulls our sensitivity to certain reasons, this kind of manic fixation seems like an obvious corrective. Consider responsibility practices: when we overlook reasons not to treat people in certain ways and they remonstrate with us, they focus our attention narrowly on those reasons by describing our error to us in emotionally charged tones. This forced attention to the reasons we elided over is what (re-)familiarises us with them and corrects our sense of the normative landscape for the future.⁴² If we are fortunate, and mania fixes our attention on the value of something that we learned to elide over, we benefit from this reacquaintance. The volume of content delivered by manic flight of ideas, or pressure of thought, means that we have a reasonably good chance of catching something we typically elide over, even if we cannot delve as deeply when each value that holds our attention is quickly replaced. It might be objected that this volume cuts both ways and could just as easily focus our attention on values we are familiar with, and which we already overvalue. The values we learn of in our work become mystical and romantic to us. I grant that this is possible but offer two

⁴² We discuss responsibility practices in Chapter 6.

responses: the first is that we do not need the manic sense of things to only familiarise us with new values to be useful, and the second is that manic fixation on the reasons we already overvalue is not as harmful as acquainting ourselves with new values is beneficial. On the first point, the moral edification that mania can provide obtains as long as we are acquainted with unfamiliar values. If we spend some time while manic obsessing over familiar values, this is time when we are not gaining the same kind of moral edification, but not necessarily time when we are backsliding. Importantly, the experience of mania is very much tied up with experiences of transcending the ordinary or mundane, and so it is likely that the novel is what is marked for salience by these experiences, and then intensely interrogated for their profound normative properties. On the second point, even if manic fixation on reasons that one overweighs causes us to romanticise these values more in non-manic life, this is not symmetrically harmful to the benefits of being sensitised to new reasons. If the agent who elides over reasons to do with clean air does so because, as in Berman's examples, their professional training provides them with a limited thick evaluative vocabulary, better articulating values in that vocabulary does not obviously make one *more* insensitive than one already was. Acquaintance with a new value creates greater opportunity to respond to reasons than deeper understanding of familiar values costs, if the latter costs any at all, which is not certain.

Fixation can also serve to familiarise us with the values we are eliding over. Focused attention on the reasons that were dismissed or the value that was disregarded is how we come to understand, and eventually name, the values that sit in our hermeneutic gaps. Whilst in Fricker's examples, hermeneutic gaps get filled by a concept being shared or emerging in discussion, (2007, 149-50), a great many concepts are developed first by an individual reflecting on their experiences. The kind of narrow focus on normative properties that mania occasions familiarise us intensely with them and offers us an opportunity to fill these gaps. We will likely do this in a messy and imperfect way, but as I will argue in the next section, there are advantages distinct to the manic nature of this process. Focused attention and deliberation is useful even in the case of ecological values. That the value is difficult to grasp does not mean it is impossible, and more time spent reflecting, with attention to the minute facets, is likely to be helpful. If these are genuinely values grounded in something on the fringes of our ken, then there is likely to be a ceiling to how well even this close attention will allow us to grasp the value. Yet, there is still an advantage to better grasping rather than failing to grasp at all. The sense of the mystic or cosmic

might be useful for getting us to the inhumanity of the natural world when we, with Hursthouse, are seeking the proper orientation towards nature.

3.4.3 Sum Up

In this section I have argued that there are families of reasons we systemically under-weigh or elide over because we are blind to the values which give rise to them. There are a number of ways, either to do with us or to do with the reasons themselves, that make these reasons opaque to us. I have argued that the manic structure of agency is useful specifically in that it focuses our attention closely and minutely on these values, familiarising us with them to a greater degree than would otherwise be likely. This kind of attentive reflection is useful for better appreciating these values in the case of mania, just as it is in the case of non-manic reflection. The manic restructuring of agency is one way, therefore, of addressing this difficulty. Manic agency is adequate, then, to solve distinctive problems for us as limited agents. It remains possible that it is one of many adequate methods, and so the next section will address what it is that is distinctive of mania that makes it uncommonly useful for this task.

§3.5 Thick Fluency

What resources does the manic agent enjoy that allows them to better grasp esoteric values that are not enjoyed by non-manic agents. These are the resources that are deployed during the period of intense focus discussed previously, that allow the agent to acquaint themselves with these values better than a neurotypical agent simply meditating on the matter is likely to. In this section I argue that various features of mania separate from those involved in the restructuring involved produce fluency in thick evaluative concepts that make agents better able to articulate values that sit in hermeneutical gaps or are on the fringes of the typical human conceptual range. My assumption here is that the ability to articulate is a reasonable proxy for the extent to which we grasp or understand these values, and thus a proxy for when we do enjoy moral knowledge. This is not to suggest that we do not enjoy moral knowledge when we cannot articulate what we know, but that we likely have limited or impaired understanding of the values in question in these cases. Certainly, some things we know without being

able to articulate, and this will be true for the manic person who finds that “it is only with a recognition of the total inadequacy of language that insight begins” (Lodge 2021, 8). We might grasp these values without being able to articulate them. Nonetheless, a greater ability to articulate values is sufficient to demonstrate a better grasp, and thick fluency (I shall argue) will confer a greater ability to articulate values.

The features of mania I focus on here are those features involved in manic patterns of association and inference which are opaque (or seem disorganised) to outside perspectives. I will draw on Sass and Pienkos’ notable (2015) study on linguistic behaviour in mania. Manic people speak quickly, speak a lot, and tend to adopt a playful approach to assonance and double-meaning which remains largely opaque to non-manic hearers. Pressure of thought and flight of ideas, between them, refer to the increased volume of thought that manic people tend to experience, but flight of ideas goes further, and refers to the seemingly attenuated connections between thoughts. These connections are often marked out by *clang associations*, associations which turn on the phonetic rather than semantic content of words (483, 486). The incongruity marked between the associated thoughts is marked from the outside, as they are typically all too plain to the manic agent themselves. Lake’s analysis of a patient whose flight of ideas was initially mistaken for thought blocking (where a train of thought stops dead, and a new one might pick up, very typical of disorganised schizophrenia) was later judged to be an instance of manic flight of ideas, obscured by the gap between the patient’s own experience and what is observable. Lake writes:

By the patient’s report, there are connections to all his thoughts, but an observer is oblivious to the connections because only about 20% of his thoughts are verbalized, and none have any relationship to the subject of “mom”. During euthymia, distraction to the keys falling to the floor would usually be blocked at the filter or given such low priority that the interview subject of “mom” would continue appropriately and uninterrupted.

(Lake 2008, 114, internal citations removed)

The thoughts were triggered by the interviewer’s key falling to the floor, and were; the key of life, the Nile river; Egypt, the Pyramids and their magnificence, the desert (from Egypt), hotness and dryness, that he wanted a glass of water. The scope of these thoughts encompasses all of human history and prehistory, enjoying the cosmic qualities common in manic thought, and drew the agent’s attention to the specifically normative properties of the pyramids (their magnificence being one of only two thoughts

that were verbalised).⁴³ Even after the connections were made explicit by the patient, Lake considers them “loosely associated” (ibid). The incongruity of the verbalised thoughts is resolved by understanding the connections apparent to the agent, however these connections remain loose to a degree worth remarking upon. The connectedness threshold for association seems to be somewhat relaxed in manic thought, compared to non-manic.

Manic thought involves high volume, frequent associations made on the grounds of things other than meaning or context, and an overall relaxed standard of connectedness required for association. Whilst this can limit the manic agent’s understandability, it tends to occur in a playful key that *strains* understandability without rejecting it wholesale, (which more seems to be a feature of schizophrenic speech, Sass & Pienkos 2015, 489). These are the resources that I argue produce *thick fluency*. Manic agents are, insofar as this is the way trains of thought unfold, particularly well equipped to narrate thick evaluative descriptions of their experience and the world that pick out, however imperfectly or imprecisely, values that are missed by more conventional ways of evaluating. The manic agent’s attention is focused on particular normative properties of things, in a way that does not hedge out other normative properties but does seem to pre-empt them, and so the evaluative component is secure. The descriptions that make this a *thick* evaluation are produced by the manic agent’s esoteric use of language, expressing in a mediated way the distinctive trains of thought. The relaxed standards for association, the ability to import associative powers from both sides of a double-meaning, and even the volume of thoughts that can be deployed as part of this narrating process.

To grasp values that sit in hermeneutic gaps requires that we go beyond our existing linguistic and hermeneutic resources. Ecological values require the same to even loosely grasp them, as the patterns of usual cognition are what are inadequate. Consider the example given above. When the concept of ‘sexual harassment’ was being developed, and the name eventually coined, what was needed was a concept that would pick out the right features of a situation, scenario, or behaviour, tie them together

⁴³ One might worry that Lake’s patient was experiencing something more like blocking, but was able to reverse engineer a set of connections to rationalise the stimulus and utterances together, in much the same way agents seem to rationalise behaviour after the fact. If this is so, it remains the case that the patient did this very quickly and produced a set of disparate associations. The ability to quickly weave dense networks of meaning with rich normative content remains present and valuable, even if it is retrospective.

into a bundle that could be referred to, (the *differentia*) and would convey that there was something terribly wrong with the behaviour that composed it. How does the manic pattern of thought do the same work? Firstly, hermeneutical gaps or ecological values are hard to grasp because we are not readily able to pick out what are and are not the features relevant to the value. To develop an adequate concept, we need to know the *differentia*.⁴⁴ Sexual harassment picks out patterns of behaviour, power relations, and social scripts, but excludes the gender of the participants. Sexual harassment is *typically* done by men to women but is not *essentially* so.

Manic description can produce patterns of inclusion and exclusion that are not as readily available to us when we adhere to typical rules of language and thought, due to the way in which manic association trades on obscure or phonetic associations and enjoys a relaxed standard of association. Trading in double meanings is often a sort of equivocation, but in manic speech (as in poetry) it also allows us to import meaning from both sides. Descriptors can then deploy larger amounts of meaning that capture features of the situation or phenomenon in a wider array of patterns. As with any norms, the set standards confine the range of patterns of inclusion and exclusion, and more relaxed norms permit a wider range of configurations. The manic person can then place ‘keys falling off a table’, ‘the aesthetic value in architecture’, and ‘their own thirst’, in the same network of meaning as grand or cosmic musings on the origins of life and civilisation. These configurations are not valuable in and of themselves, but they offer a wider vocabulary with which to describe situations, phenomena, objects of attention, or whatever else might instantiate an esoteric value. Mania offers a wider range of *differentia* upon which to build new concepts.⁴⁵ The sheer volume of manic thought plays a secondary role, yet remains important as simply having more bites at the apple makes it more likely one will successfully grasp the concept. One can run through more configurations of related ideas, importing neutral-

⁴⁴ One way of conveying that something is wrong would be for this *differentia* to be a species of the conceptual genus BAD. This species-genus model has received criticism for being Separationism about thick evaluation by another name (Kirchin 2017, 53). I do not weigh in here on debates about the structure of thick concepts any further than to say that, if manic agents are becoming acquainted with new values, both this differentiation and normative charge must be conveyed. These are necessary, but perhaps not sufficient, and whether they are in fact understandable in isolation from each other does not matter for my argument.

⁴⁵ I have in mind Kirchin’s (2013) suggestion that thick concepts have, at their core, thick descriptions (72). Whilst this account required and has received further development (Kirchin 2017), I require only the germ of the account to demonstrate the utility of thick fluency. We improve the state of our moral knowledge, especially at the periphery, when we have a description that lets one grasp an esoteric value better than they initially did, even if it falls short of the structural requirements of a discreet concept.

descriptive and thickly normative content.⁴⁶ This is evaluative as the manic restructuring of agency focuses attention intently and inflexibly on apparently urgent reasons or pre-emptive values. (Recall, it was the magnificence of the pyramids that got far enough to be verbalised by Lake's patient). These descriptions must be thick in order to do useful work in expanding our moral knowledge. If the manic agent were only quick with thin evaluations, adjudicating everything their attention landed on as good or bad, they would be opinionated rather than fluent, and would not enjoy *understanding* of esoteric values even when they landed on the right object and right evaluation.

These descriptions do not need to be particularly accessible to people other than the manic agent in order to do their work, (which is just as well, because they likely will not be). A description that highlights the right things, excludes the right things, and brings the right normative elements to bear is sufficient for the agent themselves to enjoy greater moral knowledge. One can grasp esoteric values better if one has a description that highlights the right pattern of features and with the right evaluative content distributed across the pattern. The agent has corrected themselves, bridged a hermeneutic gap, or grasped an obscure ecological value. It remains possible, however, that they will be able to share this description or concept if they continue to use it often enough, others might pick it up from context, reflect on the esoteric value themselves, and come to a similarly improved understanding. However, in the difficult cases, there is likely to still be a limit to the degree of understanding. One can get only so firm a grasp on ecological values, after all. In such cases, the ability to weave complexes of meaning, rich with normative content, gives the manic agent at least the resources to 'talk around' a value they are struggling to articulate. If we cannot speak exactly to a value or idea we are aiming at, we can still demonstrate some understanding by describing the values or concepts in the vicinity of the idea, that share logical space with it, and which make a degree of sense *inter alia* with it. What matters here is that the manic agent has an avenue to greater moral knowledge due to particular features of the manic structure of agency and its attendant patterns of thought. There is, consequently, a particular kind of moral epistemology at which this kind of agency excels, and it is an excellence worth pursuing.

⁴⁶ Often this will not produce new or interesting moral knowledge, if the object of attention is either familiar or unimportant. What matters, however, is that this configuration of mental capacities is particularly well tuned to capturing esoteric values when the agent does focus on them.

§3.6 Modes of Agency

Manic agency tends to produce fluency in thick evaluative descriptions of the objects of attention, which acquaint agents with values they would otherwise be prone to systemically under-weigh. If this acquaintance was fleeting, however, there would be limited value in this structure of agency. Indeed, we have reason to worry that it might be. The manic sense of things involves a sense of potent clarity, but this might be symmetrical with the clarity that nihilistic depressions produce. Do we have a reason to believe that the manic sense of things is more resistant than more normal outlooks to the corrosive influence of futility thinking? Depressions notwithstanding, if we cannot take something of the manic sense with us after mania remits, it will not leave us better edified for the future. Lodge frames the dilemma in terms of the "echo of what now seems like a distant land" (2020, 7), suggesting that it is already mostly lost. This section will argue that a way of retaining the manic sense of things is available. The manic sense of things can be preserved as part of our curated *inventory of agency*, (Nguyen 2019, 2020).

Nguyen argues that games are works of art conveyed through the medium of agency (rather than through 'focused light and sound', or 'oils on canvas'). He argues that, when we play games, we *submerge* ourselves in layers of agency (2019), or in modes of agency (2020). Modes of agency are "focused ways of being an agent" (2020, 79). Modes of agency are individuated by the sets of facts that count as reasons per that mode, and the range of responses those reasons might call for per the mode, (see §2.3). When we play chess, our mode of agency highlights a particular range of facts about the positions of different pieces on a board and their institutional (but not spatial) relations to each other as reasons, and offers a confined range of responses that these reasons can call for. If your Queen controls her side of the board, this might call for my Knight to K5, or an attack by my Bishop to force you to redeploy her, but it does not call for me to move her to another square (she is your piece). A game such as *Risk* involves more variables, allowing more states of play, meaning that a wider range of facts can be reasons per the game's mode. Similarly, *Risk* permits a wider range of responses, both in their being more ways pieces can be placed on the board, but also because alliances, ententes, and betrayals are key elements of play. On Nguyen's analysis, "games don't just simply describe the outlines of such an agency; they plunge the player into it, exposing the player to that form of agency from the inside." (80)

Crucially, on this account, basically competent agents can switch between modes of agency with reasonable fluency. The agential modes involved in contributing to (but not dominating) conversation, and managing a first aid situation, are drastically different, but a competent first aider, medical student, nurse, etc is someone who can switch from the former to the latter at the drop of a hat when an emergency arises.

This is the model according to which we can retain something of the manic sense of things. Just as games familiarise us with modes of agency, manic episodes familiarise us with modes of agency which focus intensely on the particular, perhaps obscure, values that are the objects of our attention. The urgency of mania focuses us on particular reasons, as any mode would do, and the thick fluency mania occasions allows us to focus on reasons we might otherwise struggle to grasp. On this side of the coin, mania produces a mode of agency that is sensitive to esoteric values. How then do these modes of agency distinguish themselves in terms of the responses available?⁴⁷ The urgency of mania makes a difference. Whatever reason we are fixated on is surrounded by other reasons that moderate what can be called for, all things considered, but given our sense of urgency, those surrounding reasons 'need not be weighed up'. The manic agent finds themselves with more latitude in response to reasons given the priority enjoyed by their urgent reason. However, given the manic agent experiences this consideration as rather more like a requirement, strict failure conditions seem to hold, in another sense narrowing the latitude the agent has. These changes make manic action more like Chess or *Risk*, respectively in their different senses. This is sufficient to let us start to individuate modes of manic agency, however it will help to have more specific descriptions. Whilst manic urgency might broaden or narrow what responses are available to us, nothing in the analysis so far tells us what kinds of practical actions might belong to these modes (in the way that 'moving pieces' belongs to both Chess and *Risk*, and 'making and breaking alliances' belongs to *Risk*). Considering the phenomenology we have explored in the previous chapter, we can recognise sorts of response that seem to flow very easily in manic states. Manic modes of agency seem to privilege *appreciation* as appropriate response. Lodge describes the drive to 'speak to the sheer being of things' (2020, 7). Some reasons do not call for practical action in the same way that others do. There are some injustices that cannot be rectified by any action available to us, but we ought

⁴⁷ This is important, as if mania only sensitised us to new reasons but did not provide new ways of responding, it would not be a mode that we add to an inventory as much as it would be a (potentially short-lived) change in our normative outlook (see Chapter 1). The objection that the manic sense of things is fleeting would remain unanswered.

ruē nonetheless. The virtue of *wonder*, as Hursthouse considered it (2007), is chiefly to do with appreciating nature, more than it is to do with acting upon or towards it. The reports of manic experience we have explored tend to include excitement and a sense of profundity or cosmic significance as prominent features relevant to action. We should, therefore, recognise a broad family of ways of *appreciating* value to be the responses to reasons.⁴⁸ Manic modes of agency can sensitise us to reasons and values we otherwise struggle to get a grip on, and offer a range of primarily appreciative responses that those values can call out for. Appreciation might be a private mental exercise or a specific physical action (like waxing lyrical about something), but appreciation can also be multiply realisable, achieved by any combination of other actions down to the manic agent to decide on.

Nguyen argues that games support our autonomy by familiarising us with particular modes of agency that might be useful to us in our broader agential lives. Whilst highly restrictive in the sets of reasons they admit of, such modes of agency intimately familiarise us with those reasons and the kinds of response they call for. He summarises his argument as follows:

Wide exposure to games can enhance the autonomy of agents by making them aware of alternative modes of agency. By communicating agencies, we can enlarge and enrich each other's autonomy. We can help to broaden each other's knowledge of agencies, help to develop our capacities to switch between modes of agency, and our ability to find the right mode of agency. Games, as a medium for communicating agencies, can help us to develop our autonomy as a collaborative project. We can think of new forms of agency, write them down in games, and share them with one another. We can refine modes of agency and store them for future generations.

(Nguyen 2020, 78)

Manic modes of agency are products of the re-structuring of agency discussed in the previous chapter. It is part of the existential reality of living with BoRD that these modes of agency will impose themselves. Manic episodes are the period during which agents are 'plunged into' a manic mode of agency. This is how mania familiarises us in much the same way games do, with the reasons that mode is sensitive to and the kinds of response (appreciation) they call for. Games are artifacts that communicate these

⁴⁸ I overstate the distinction between practical action and appreciation. Certainly, there are some cases in which our appreciation of some value is a passive experience, anger at injustices done to us for instance, but we can also make efforts to actively appreciate values. This might be simply a mental exercise, but it can also be something we do via physical action. We might privately muse on the glory of being young and in love by focusing our attention on the matter, or we might cry *εὐοῖ* and dance about it.

agencies between people. Manic modes are not communicated between people and are not communicated through artifacts. How then do we retain the manic sense of things and add it to our curated inventory of agencies? We hold onto the descriptions produced by our thick fluency, which pick out features of the world, patterns of facts, and cast them as calling for some sort of normatively rich appreciation. These thick descriptions are the artifacts that we store and curate, adjusting with the benefit of non-manic reflection, and going back to during depressions, when we need something to slow the rate of decay. Our inventories of agency are supplemented by pictures of the world that tell us, even if we cannot feel it in that moment, that there is great value and meaning to be found in the changing of the seasons, or the movement of the stars, or the excitable games we played as children. In Chapter 1 we discussed the way in which two motivational systems train and discipline each other. The thick descriptions we store as parts of our inventories function in much the same way judgements do, highlighting facts as normatively salient and valanced, with independent motivational power even when our emotional experiences do not paint the same normative picture. Through the same process by which depressions wear away our sense of meaning, the preserved manic sense of things applies an opposing pressure (both during depressions and outwith them) sensitising us to the value of things we were aware of during mania.

3.7 Advantages of Manic Insight

Now that all our pieces are arranged, what do they amount to? Above, I outlined the argument from depression and the argument from moral edification. I will reiterate them both, now that the resources are in place to demonstrate their persuasiveness.

3.7.1 Edification

The edification argument holds that there is moral knowledge available to us that we are more likely to enjoy access to during mania than during non-manic states. The argument runs as follows:

P1) We systemically under-weigh certain (families of) reasons

P2) Mania fixes our attention narrowly and deeply on certain values.

P3) This deep and narrow attention can correct for under-weighing.

C1) Mania can provide a resource to correct for systemic under-weighing.

The moral knowledge available is knowledge of the values we systematically under-weigh. The restructuring of mania fixes our attention narrowly and, combined with thick fluency, allows us to better grasp those esoteric values, letting us appreciate them deeply, and not under weigh them. Combined with Nguyen’s analysis, we gain the following premise and second conclusion:

P4) This deep and narrow focus can be stored in our inventories of agency

C2) Mania’s corrective resource can be useful for the future.

Thick descriptions of (parts of) the world that we compose during manic flurries are descriptions we can hold on to and trade in at later dates. These descriptions are the artifacts we use to store agencies in our private inventories. Thus, the edifying and corrective effect of the manic sense of things can be relied upon in the future.

On the reasons-responsiveness conception of autonomy I have deployed, we excel better as agents for having these modes as elements in our curated inventories of agency. With these resources at hand, we are better equipped to respond to esoteric values when we encounter them, to be able to fill in hermeneutic gaps in our moral schema, and to pull ourselves out of well habituated but limiting modes of agency that blind us to some values. The particular kinds of response that manic modes provide are appreciative responses. The manic agent’s abnormal normative experience gives rise to rarefied moral knowledge. This observation has been made before. Jamison remarks on the “depth, fire, and understanding” that the artistic imagination benefits from, and the “different kind of insight, compassion, and expression of the human condition” that manically depressive artists contribute (1994 102). McNamara & DuBrul remark on the namesake of the ‘Icarus Project’:

the story of Icarus’ wings has served to remind us that sometimes the most incredible of gifts can also be the most dangerous. With our double-edged blessings we have the ability to fly to places of great vision and creativity, but like the mythical boy Icarus, we also have the potential to fly dangerously close to the sun—into realms of delusion and psychosis—and crash in a blaze of fire and confusion.

(McNamara & DuBrul 2007, 1)

This chapter joins these moves in highlighting the moral-epistemic advantages one can gain from relating in the correct way to manic experiences, and contributes a specific account of the way in which mania confers these moral-epistemic advantages. The account of the manic restructuring of agency offered in the last chapter, as well as the thick fluency explored here, explain how the manic agent gains this spiritual leg-up. Episodic mania is distinctive in that the intensity of feeling involved allows us to feel, to an extent, what it is like to be the agent who values in this way. As it typically remits and recurs, the manic agent is uncommonly placed, able to try on outlooks like clothes. It is a transformative change one comes back from (compare, Paul 2014).

Being able to recognise and respond to values and reasons in more cases is an advantage to any being for whom the description agent (understood in terms of reasons-responsiveness) applies, and so these advantages contribute to the flourishing of manic agents qua agents. We are, however, many things besides agents, and it is the way in which the manic sense of things attunes one to meaning in life that perhaps makes the greater contribution to a manic person's flourishing. Wolf argues that lives are meaningful to the extent that we are "actively, and to some extent successfully, engaged in projects (understanding this term broadly) that not just seem to have positive value, but that really do have it" (2015, 96). Projects here must refer to something that ties together reasonable portions of life, contributing to a unity (95f4), but beyond that, can range widely across relationships, activities, involvements, "or even just from communing with or actively appreciating what is there to be appreciated" (112). It is this last way of finding meaning that mania seems especially suited to. The greater freedom in mental grammar that manic thought involves permits us to grasp esoteric or ecological values. This admits a wider range of meanings into life, but importantly, it admits meanings that are less anthropocentric and more sensitive to the genuinely *other*. Recognising the world as something to which has value in its own terms, to which we can be responsible, and with which we can be in community, introduces meanings to life that leave us less normatively alone. To the extent that received value systems are inclined to consumption, disposal, and eventually, destruction (the anxiety that animates the environmental literature above), the manic sense of things is particularly well equipped to help us reach the corrective. Insofar as there are more things to appreciate, people who take the manic sense of things seriously have wider latitude for weaving meaning into their lives. More projects can satisfy the conditions Wolf sets out and play the meaning-conferring role, but more importantly, more elements of existing projects might be able to do this. But what is perhaps most

valuable is that the manic sense of things is what can most effectively force the normal human perspective on the world open a little wider, highlighting that we are not normatively alone.

3.7.2 Argument from Depression

The argument from depression matters specifically for agents living with BoRD as they are likely to experience depressions as well as manic episodes. If relating to manic experiences in a particular way could serve to make depressions more survivable, then agents who experience both have good reasons to relate to their mania in that way. The argument is very much as above, admitting of only minor changes. Rather than systemically under-weighting certain (families of) reasons being a premise supported by an account of hermeneutic gaps or esoteric values, something like it is supported with reference to the futility thinking that many agents with BoRD endure. Another difference is that the tenses are slightly different. The under-weighting is not habitual, but localised to depressions, in the same way the mania is localised in P2 above. Finally, we do not gain a satisfying conclusion at C1, as the problem with futility thinking is it will mislead us while it obtains, and it tends to shut out other outlooks. Making these changes, the argument looks like this:

P1) Depression will routinely blind us to certain (families of) reasons (Futility Thinking).

P2) Mania fixes our attention narrowly and deeply on certain values.

P3) This deep and narrow attention can correct for under-weighting.

C1) Mania amounts to a resource to correct for systemic under-weighting.

P4) This deep and narrow focus can be stored in our inventories of agency and accessed at later dates, including during depressions.

C2) Mania's corrective resource can be used to correct for futility thinking.

The first premise is the core of our analysis from Chapter 1. P2, P3, and C1 are as above. P4 is a more specified form of the analysis built on Nguyen's account. C2 is the conclusion of the argument from depression, and what seems to be indicated by Plato's implicit argument about Dionysian madness. Manic episodes, if properly related to after the fact, offer a medicine journey, letting agents return with elixirs that can fortify them against the distinctive challenges they are more likely to face. Having either a stronger sense of meaning, or more things in life that seem meaningful, is unlikely to prevent many

depressive episodes from emerging, however it is more likely to slow the rate at which futility thinking eats away at one's sense of meaning or one's moral knowledge. Depressions can be made more survivable if they are limited in how deep they can become, or the rate of decay is slowed, or if there are more values to serve as handholds with which agents can try to work themselves out of the malaise. This will not be useful against every sort of depression, there are far too many ways to be depressed for a panacea. But for those depressions which are harmful to our agency, happiness, or flourishing specifically because of the way they limit our moral sight, taking the manic sense seriously can be a valuable exercise in self-care.

3.8 Conclusion

The previous chapter developed and argued for a model of the restructured agency occasioned by mania. The pervasive urgency that this agency involves presents obstacles to reasons-response for manic agents and therefore threatens manic agents' autonomy, insofar as autonomy is built on reliably reasons-responsive agency. Manic agency remains sufficiently capable of tracking and trading in reasons, however, even if the relative weighing will be erratic. The danger was of impaired, but not arrested, autonomy. This chapter demonstrates that there is a specialised task for which the modes of agency occasioned by mania are specifically suited to. The limitations acknowledged in the previous chapter are not sufficient to instruct us on whether manic agency can be excellent in reasons-responsiveness until we have inquired as to whether there is a specialist role it might serve, despite those limits. This chapter has argued that such specialist work exists and that manic agency is well suited to it.

Manic agents might struggle to respond to reasons in balance with each other, due to the structural effects of urgency, but are especially well equipped to learn about and respond to obscure or esoteric values. The kind of response that manic agents specialise in is appreciation, which might variously involve musing on, celebrating, or communing with value in a way that invites meaning into life. Thus, whilst manic agency is not an excellent generalist mode of agency, it can be an excellent specialist

mode. Understanding autonomy as an ideal in terms of the excellent functioning of our agential capacities, manic modes of agency can contribute to our flourishing as agents by carrying out this specialist work. Insofar as I am interested in the kind of agency available to agents living with BoRD, this specialist work would seem to play a greater role in the flourishing of the kind(s) of agent someone with BoRD can become. Whilst different inventories of agency are likely to be more suitable to different kinds of agent, we should expect the specialised mode to play different roles for different agents. The fact that the specialised manic mode is so readily available to agents with BoRD means that such agents can more readily become, and perhaps have better reason to pursue, kinds of agency in which this mode plays a greater role. Crafting the kind of agency we will have must be done within the confines of our existential reality, and can be done more easily working with the grain. This does not mean that agents ought to only pursue the kinds of agency that come most naturally. All it means is that this is the most readily accessible sort of excellence, and so they have good reason to craft for themselves the kind of agency with which they are more likely to flourish. There is still something noble about fighting against the grain for the sake of a less accessible goal. The perversity only arises if the less accessible value is 'normalcy' of a sort that is hedged out or made inaccessible by one's existential reality.

A final consideration is what kind of recognition is owed to this mode of agency. Often it will be hard to recognise, as complex and arbitrary seeming networks of action might be what add up to appreciating or celebrating some esoteric value. If the answer to *why* is confusing to non-manic people, the kinds of social pressures Lodge highlights are likely to arise. However, if we understand manic agency not as defective functioning, nor as the practical expression of disorder, but as a specialised kind of agency, we can more easily recognise that these complex and confusing performances are expressions of some kind of agency. Pathologizing stances withhold this recognition, and so understanding manic people's action and speech requires an agential stance, in order to recognise this agency. A simple way to express this is to accept manic speech, with its confusing answers to the *why* question, as moral testimony. The manic person who is waxing lyrical about the brilliance of Jamison's paperback pun, or how enraged they are by atrocities from the Bronze Age, is offering normative testimony. Recognition does not require that one weigh heavily, or even make time to listen to, testimony. Indeed, this would place an unreasonable burden on strangers a manic person meets at a party.⁴⁹ Further, constantly asking a manic person to repeat themselves, to explain associations, or run through an idea more slowly, will only

⁴⁹ Jamison offers an illustrative anecdote, (1995, 70-1).

infuriate both parties. Recognition can be conferred by accepting that testimony was delivered even if it was not caught or understood, as long as it is read as testimony and not symptom.

I take the ideal of autonomy to involve the excellent functioning of agency, whilst admitting that 'agent' is an underspecified description and that there are many kinds of agency that, when properly specified, reveal quite distinct excellence or flourishing conditions. Over the last three chapters what has emerged is a picture of the kinds of agency that someone living with BoRD will find themselves with at different times, and accounts of the differences we find in the excellence of those kinds of agency. Contrary to the presumption that remains influential, the fact of madness does not impair autonomy in the sense of making the ideal less accessible. Rather, madness that affects agency will often change what kind of agency is at issue, thereby changing what the ideal of autonomy, local to that kind of agency, looks like. There is specific work that can be done excellently by manic modes of agency, which supplements the work done by more neurotypical agencies. With a picture in place that tells us what excellent functioning of agency looks like in this corner of the warren, we must now understand how that agency can be scaffolded across time. BoRD is marked by recurring affective shifts, between the kinds of agency we have examined so far. Bipolar agency extends through and across different structures of agency and so being the kind of agent one is will often require supportive scaffolding. The next two chapters will examine the sorts of scaffolding that support excellent functioning of bipolar agency, in light of the picture we currently have in place.

Chapter 4 – Self-Narrative and Self-Trust in Bipolar Disorder

§4.1 - Introduction

In the previous chapter we considered the dilemma bipolar agents face in the aftermath of episodes of mania: whether or not to take the manic sense of things seriously, as offering moral testimony. In this chapter we return to the reasons given against dismissing the manic sense of things. We were concerned in §3.2.3 that by dismissing the manic sense of things we might become alienated from our own normative outlooks in a way that arrests our agency. The argument was that explaining our sense of what was valuable with reference to our mood, rather than with reference to our capacities to respond to value, creates an undercutting defeater to any moral knowledge, irrespective of which mood state we are in. Such pervasive undercutting of our moral knowledge would erode our self-trust to such a degree that our autonomy would be severely threatened. Avoiding this alienation is, therefore, a reason to take the manic sense of things seriously.

It might be argued, however, that this does not adequately meet the challenge to self-trust that episodic affective shifts give rise to. Episodes of mania or depression disrupt the self-narratives that agents rely on to secure self-trust. Rather than mania or depression being threats to self-trust themselves, it is the fact that they episodically emerge and remit. In this chapter I argue that self-trust in the face of bipolar affective shifts is more likely to be secured by self-narratives that involve the structural features of medicine journey narratives. The next section will outline the attitudes involved in self-trust and the role they must play if we are to enjoy tolerable degrees of autonomy. §4.3 will highlight the specific threat that BoRD creates for self-trust, which will make clearer that the problem we discussed in §3.2.3 still looms. §4.4 will establish our desiderata for an adequate self-narrative. §4.5 will discuss two bipolar narratives that have been proposed as specifically suitable to make sense of the existential reality of BoRD whilst also securing self-trust. This section will examine these narratives against our desiderata and highlight the structural features that satisfy our desiderata but note that neither is wholly adequate as it has been offered in the literature. §4.6 will offer the medicine journey narrative as a structure that preserves the desirable features of previous narratives and argue that it more fully satisfies our desiderata. §4.7 will outline the consequences such a self-narrative can have for the autonomy of bipolar agents, particularly in the self-determination dimension.

§4.2 - Self-Trust

In the aftermath of mood episodes, we might be tempted to dismiss the normative outlooks we held during those episodes. Remissions allow the space to critically reflect on the normative pictures of the world found in manic or depressive episodes. The fact that we now enjoy such a distinctly different normative outlook is likely to injure our confidence in our outlook. When we disagree with ourselves in this way, at least one of our two contrary views must be mistaken. The mistake is evidence that we are not as reliable as we might hope. This creates pressure to adjust our confidence in our own assessment downwards.⁵⁰ Different forces can drive this atrophy in self-trust, and self-*distrust* will find expression in different attitudes. McLeod and Sherwin describe it as follows:

The primary influence on self-distrust that arises through inductive reasoning about past failures will probably be a belief, whereas the primary influence on self-distrust that exists because of subtle attempts by others to undermine self-trust will probably be a feeling about our own incompetency.

(McLeod & Sherwin, 2000, 265)

§4.2.1 - Self-Trust and Autonomy

McLeod and Sherwin distinguish three kinds of self-trust. The first is the trust we have in our procedural judgement; that our values will be reflected in our practical deliberation. The second is trust in our tendency to follow through, that we can depend on our (future) selves to carry out commitments. The final kind is our trust in the appropriateness of our own judgement: that we are in fact 'getting it right' when we take something to be valuable in the way we do, (2000, 263-4). To different degrees, each of these are injured by the kind of doubt that BoRD gives rise to. We might doubt our procedural judgement when we think about our depressive tendency to catastrophise and discount the positive, or when we think about our manic tendencies to leap thoughtlessly into rash action. The second type of trust is undermined by the knowledge that affective shifts will disrupt our plans, as depression causes us to give up on projects that are near completion, and mania causes us to overthrow long standing commitments for the sake of exciting new projects. The temptation to understand our moods as

⁵⁰ Note that we do not need to grant the evidentiary view of disagreement for this argument to hold (Christensen 2009, 2010). Whilst we can describe this as a diachronic disagreement, the higher order evidence is grounded in the functioning of our own epistemic capacities, rather than in some independent epistemic agent who acts as a check on us.

something alien to us, or as essentially incompetent, undermines the final kind. Living with BoRD involves the complex task of forming an epistemically well-founded view of one's own capacities and competence, whilst also maintaining the kind of self-trust that is crucial to enjoy tolerable degrees of autonomy.⁵¹

Some degree of self-trust is necessary for autonomy. If we do not trust our judgement, either procedurally or substantively, we are less likely to rely on it. We are more likely to depend on other agents or external factors to guide what we should do in this or that situation. This kind of dependence injures the degree to which we are self-directing, rather than being directed by others, (McLeod & Sherwin 2000, 264). If we don't trust ourselves, we will be reticent to make self-determinations. Distrusting our (procedural or substantive) judgement means that we will lack the confidence to close deliberation by committing to some course of action. It is precisely when it seems to an agent that their best all-things-considered reasons are to ϕ that it is important they have the trust in their own sense of these things. The possibility that some other factor which we have not noticed is out there, which makes it the case that one should not- ϕ , is an ever-present stalking horse. It can't be proven false, and so is the degree of doubt we must live with. Self-trust involves the attitude that one's own agential capacities are reliable enough to justify closing deliberation. Self-distrust makes us hesitant to exercise our autonomy by deploying agential capacities in self-determining ways, and so shrinks the footprint of our agency, upon which autonomy is to be founded.

Above this minimal threshold, self-trust continues to facilitate agents' autonomy. Bortolotti has recently argued that beliefs which foster self-trust, even when they are epistemically irrational, serve highly useful roles in supporting our effective agency. She notes significant (although domain specific) correlations between confidence that we are competent or likely to succeed, which exceeds what is rational, and high degrees of success and efficacy, (2020, 114). More importantly, several of the epistemic errors we make when trying to preserve self-trust seem to be self-fulfilling:

⁵¹ McLeod and Sherwin speak of judgement, but we can trust our judgement whilst doubting our sentimental motivations, as many people with BoRD do. This is qualified self-trust and qualified self-distrust, and so we should perhaps replace the term 'judgement' with 'motivational systems'. Govier (1993) had previously adopted something like this formulation, taking self-trust to consist in trust that one is well-motivated (105). Aside from this change, I keep to McLeod and Sherwin's formulation because they more clearly individuate the kinds of error we make when we engage in different kinds of self-distrust.

Despite their being ill-grounded and often inaccurate, every day confabulations contribute to our sense of ourselves as competent agents and decision makers. In addition to that, they can also help us create patterns for our judgements and choices that make our overall commitments more memorable and more meaningful to ourselves and others.

(Bortolotti 2020, 53)

Bortolotti picks up on a thread that Summers has previously explored (2017), highlighting the ways in which we confabulate plausible and plausibly justified motivations for our actions as ways of habituating sensitivity to those reasons in ourselves. The more frequently we offer our concern for animal welfare as our reason we stopped eating meat (when it was really a miniature rebellion against our parents) the more the pressure to be consistent makes it likely that we will be effectively motivated by similar reasons (to avoid unnecessary leather and minimize palm oil consumption). We confabulate in this way, he suggests, to maintain an image of ourselves as rational, virtuous, and deserving of self-trust. It is a self-fulfilling report, as maintaining such an image of our selves makes it more accurate than it otherwise would be. Self-trust supports our basic agential capacities by creating demanding expectations on us that we are driven to meet. The agential capacities upon which autonomy is built are supported by optimistic and somewhat revisionary characterisations of our motivations. Self-narratives that trade in these confabulations inspire the kinds of self-trust that are autonomy-supportive and scaffold the agency that deserves that self-trust.

To the extent that autonomy is socially constituted, self-trust facilitates our autonomy by altering our behaviour in relevant ways. Agents who trust themselves are more likely to offer their perspective of reasons in social interactions, having confidence as they do both in the value of their perspective and the likelihood of its being appropriate. Offering a perspective on reasons prompts others to take an agentic stance towards us, even in cases where others disagree with us or object to our perspectives. Indeed, disagreement can give rise to exchanges on the point which familiarise others more deeply with our perspectives. This is valuable particularly for agents whose autonomy is systemically threatened by pathologizing stances. Deeper familiarity with mad perspectives on reasons, even if they are ultimately rejected each time, makes interlocutors more likely to recognise mad people's utterances as offering a genuine perspective, and so demanding the agential recognition that is necessary for socially constituted autonomy. This kind of interaction strengthens the foundations of autonomy in another respect, insofar as interaction with other people invites useful scrutiny of our normative outlook and our plans, (Bortolotti 2020, 54). This scrutiny might injure self-trust if it is unconstructive and pervasive, but it can

also spot the problems we overlook in our own deliberations and prompt us to make more careful plans, or to interrogate our values more carefully. Revising our evaluative judgements or plans in the light of these interactions insulates us from a degree of frustration or failure that reduce our efficacy as agents and our self-trust.

§4.2.2 - Self-Trust and Self-Narrative

An agent's self-narrative is the story they use to make sense of their actions and experiences, and which they offer others to do the same from an external perspective. Our histories, possible futures, relationships, and roles in the world are all usually elements of our self-narratives. We narrate a self by putting these elements in sense-making juxtapositions, often informed by a set of received hermeneutic resources. These received resources are shared interpretive tools, models, tropes, and identities that are cultivated by epistemic communities, and might be more or less adequate to meet the interpretive needs of some person or group.⁵² We will offer different narratives in different contexts, perhaps narrating together larger or smaller portions of our life, spread across time and through different domains, as needed.

A convenient way of talking about narratives is as involving a central metaphor. The metaphor that is central to a narrative is the one that is doing the key sense-making work. We call some narratives 'quest narratives', even if the person composing that self-narrative never uses the term or thinks about their experiences as a quest. The quest analogy helps us understand this kind of narrative as goal oriented across the long-term, as involving obstacles, as achievable by a number of different paths, as something throughout with we will undergo transformative change, etc (see MacIntyre, 1985, Chapter 15). Our self-narratives will often involve a great many metaphors, and other sense-making elements, and so they might be describable as more than one kind of narrative (a journey narrative, a quest narrative, and a restitution narrative all at once). I will refer to narratives as individuated by these kinds of central metaphor, but this does not indicate hard and fast distinctions between these narratives. Rather, it allows us to draw comparisons between the structural features of these narratives that either facilitate, or do not facilitate, self-trust and autonomy.

⁵² For fuller discussions, see Fricker (2007, Chapter 7) and Rashed (2019, especially Chapters 5 and 6).

Self-narratives play important and wide-ranging roles in agent autonomy, one of which is in the establishment and maintenance of self-trust. Self-narratives establish self-trust by *capacitating* us as agents. By ascribing a capacity to an agent, we create an expectation in that agent and in their peers that they avail themselves of that capacity. McGeer and Pettit argue that this is the way in which holding agents responsible scaffolds virtuous agency, (2015, 167). I argue that self-narration capacitates us in a similar way. When we face some difficult problem, one which we are unsure we can resolve, we capacitate ourselves by casting our situation as the front of an unfolding narrative in which our success is likely. The narrative might draw in a number of similarly difficult problems we have overcome in the past, or it might draw in our history of developing certain skills, the culmination of which development would be our overcoming this problem. By casting the current problem as the latest in a line of these predicaments we represent our predicaments as calling for particular modes of agency and ourselves as being suitably equipped to respond. This is why our self-narratives can go a step beyond what an epistemically well-grounded report would say. Higher expectations, up to a point, ascribe greater capacities to us. Thus, aspirational self-narratives play a valuable role in helping us become more virtuous or effective than we are.

Self-narratives also protect self-trust once it has been achieved. When we are confronted by our limitations, we can self-narrate in ways that discount these failures as evidence of our incompetence. This tendency might quickly become maladaptive, but for agents with otherwise rather fragile self-trust, it can be more conducive to their autonomy to self-interpret with a slight rose-tint. Tekin notes that diagnostic concepts, such as major depressive disorder, can perform this role. She suggests that:

[a]fter being diagnosed, [an agent] may make better sense of, say, her increasing sadness, significant weight loss, insomnia, and suicidal thoughts exacerbated by feelings of hopelessness and despair in her early adulthood. This understanding may lead her to reassess her failure in her first job as an outcome of mental disorder, instead of, say, incompetence. ... She may, for instance, blame a failed relationship on her depression.

(Tekin 2011 365)

Diagnoses offer concepts with which we can narrate our stories in a way that places certain actions, experiences, frustrations, or failures on the predicament side of a protagonist-predicament distinction (Schroeder 2022, 219). These experiences, actions, or failures which had previously been evidence of our untrustworthiness are redeployed as narrative elements that are at least neutral with respect to our

trustworthiness. Rather than being incapable, we were in harsh circumstances. We can consider ourselves to be competent agents, simply unfortunate in where we find ourselves. This can be useful insofar as it grants agents permission to trust themselves in spite of failures, because the failure can be attributed to difficult circumstances. We can go too far in this practice, as shall be explored below (§4.4), but the practice of re-narrating life in this way is at least capable of supporting our autonomy by re-enforcing self-trust.

§4.3 - BoRD and Self Trust

We briefly noted, in §4.2.1, some straightforward ways in which BoRD can endanger self-trust. If we think of mania and depression as intruding or necessarily irrational forces, we will doubt that our motivation is procedurally sound (costing type 1 self-trust). It doesn't take much experience of the affective shifts that characterise BoRD to start building the data-set that casts doubt on an agent's ability to see through their commitments. Bipolar agents who have not developed self-management techniques that effectively scaffold their agency across time risk coming to believe that they cannot be relied upon to be effective across time (costing type 2 self-trust). If we think of mania or depression as morally incompetent moods, then we are likely to doubt that our motivations are substantively sound (costing type 3 self-trust).

There are deeper running and less direct ways in which BoRD disrupts self-trust, however. Self-narratives can fail to secure trust when successive affective shifts prevent us from relying on familiar narratives. Narratives that we had relied on previously, or which are generally accepted in our social circles, might suddenly ring false, no longer making the world or our experiences intelligible to us, and failing to present our pasts or possible futures in ways that we recognise as our own. We cannot rely on these narratives anymore because the narrative connections they involved relied on a particular normative outlook we held at the time. If we no longer hold that normative outlook, the self-narrative has become *empathically inaccessible* to us (Schectman 2001), and the narrative will ring false. This lack of fit can take various forms. We might continue to use the narratives we previously relied on to interpret our life and find ourselves frustrated when the sense-making that we expected to happen fails to come about, and we feel that we are merely impersonating ourselves, (Calhoun 2009, 201). In other cases, we might abandon those narratives and offer fresh interpretations of our lives. Recall from

Tolstoy's account of his depression (§1.3.1), that he redeployed narrative elements of his life not as objects value, but as tricks and deceptions that he has overcome. The final possibility is that we find ourselves without any self-narrative to rely on. Potter describes this as a *chaos narrative*:

in which [bipolar agents] narrate their lives as hopelessly out of control, senseless, and fragmented, with little hope of improvement. They may be able to identify reportable events (say, a costly shopping spree), and even situate those events in the context of other events (say, the consequences of the spree), but cannot identify the agentic or causal elements beyond the illness itself to piece together a picture of their lives that is not paradoxically both chaotic and deterministic.

(Potter 2013, 60)

We might even doubt that this is any narrative at all, if the connections that hold between narrative elements fail to do the essential sense-making work of narrative. It makes little difference whether we describe this as being left with chaos *instead of* narrative, or chaos *as* narrative. The agent is stuck unable to make sense of their experiences, and so cannot represent their experiences as evidence of their trustworthiness.

In manic moods, it is more likely to be humility and keen senses of our limitations that are empathically inaccessible to us. Whilst this looks superficially like the kind of biased optimism that correlates with success and efficacy, it cannot always be relied upon to provide those benefits. For all the work they do capacitating us and presenting the world at a resolution where opportunities call out to us, our self-narratives also remind us of our limitations. The way we imagine ourselves when narrating is informed by our previous successes and failures. Understanding some things as personal limitations is important for retaining self-trust, as when we run up against these limitations, we understand ourselves as competent, but not omnipotent. Without a narrative which tells us in which domains and circumstances we might expect failure, these failures are free-floating evidence of our ineffectiveness. Whilst failure will always apply a downwards normative pressure on our sense of our own capabilities, that pressure is contained when deployed in a narrative that accounts for those failures. When mania causes us to disregard our limitations, we are at risk of this kind of trust-damaging failure. Manic optimism often leads us to take on projects that are firmly beyond our abilities, and so we are likely to fail. Without an interpretive resource that lets us deploy that failure in intelligible ways that do not upset our sense of ourselves as effective and competent, our self-trust is endangered.

This is not a problem of BoRD injuring an agent's narrative capacities, but rather of their existential reality being one for which commonplace narratives are ill fitted. Our practices of self-narration draw deeply from a pool of shared interpretive resources, which have been developed by communities to serve common purposes. These received narratives are, consequently, biased towards meeting the interpretive needs of neurotypical community members. The difference in existential reality which means that people with BoRD are people with minority minds is the difference that gives rise to this problem. The shared pool of interpretive resources is adapted for a different kind of existential reality, marked by greater stability, lower degrees of disagreement between outlooks, and slower transformative changes than is expected in BoRD.⁵³

For people who already find that trying to rely on their own normative outlook leads to frustration, this kind of social pressure is likely to drive breakdowns in self-trust. Self-trust is something that must be facilitated by one's social environment, and if that environment does not provide for the interpretations that give rise to self-trust, or insists upon interpretations that are inimical to self-trust, then they are social relations that are inimical to autonomy.

§4.4 - Adequate Narratives

The previous section demonstrated that BoRD tends to injure self-trust in ways that threaten our autonomy. Some of these injuries to self-trust come in the form of disruption to the self-narratives agents typically use to establish and maintain self-trust. The other injuries to self-trust are likely to be ameliorated by adequate self-narratives that represented them as manageable challenges, or limitations to an otherwise competent and reliable kind of agency. If bipolar agents are to enjoy what autonomy is available to them, they require self-trust, and so will require adequate self-narratives. Received

⁵³ It might be suggested that this inadequacy of received narratives to meet the interpretive needs of certain minority minds amounts to a hermeneutic injustice, (Fricker 2007, 147). Whether this is the case or not turns on whether the lacuna in the community's hermeneutic resources is the result of mad people being prevented from participating in the process of collectively producing these interpretive resources. It is certainly the case that 'mad person' is a hermeneutically marginalised identity, but it does not follow from this fact that the lacuna would be filled if mad people were less hermeneutically marginalised. People living with BoRD trade in different kinds of narrative to different degrees of success. This might simply be a more difficult hermeneutic problem to solve, and so one that takes longer for any community to develop adequate resources for. Living in the time prior to those resources being developed is epistemic bad luck, but not itself an injustice (ibid, 152).

narratives tend not to be adequate. What features would an adequate narrative have, and what should we hope it would achieve? So far we have established two desiderata. The first is that it keeps a narrative the bipolar agent open to their own experience of value, which is important for the reasons explored in the previous chapter. Secondly, the narrative should help the agent establish and maintain self-trust. The remainder of this chapter will examine kinds of narrative that have been suggested as involving suitable features to do this work, and argue that the *medicine journey* offers at least a suitable, and perhaps the best, narrative structure. But first, let us add one more desideratum.

There is another sort of narration that can be dangerous for autonomy, but which is tempting when received narratives fail to be adequate. It is to revise the history that is to be narrated, to refuse to believe certain propositions that are evidently true, or to draw narrative connections in a way that could not, even in principle, be acceptable to interlocutors. BoRD does present an increased risk of revisionary, or even delusional, narration. The sense of revelation and great importance that manic episodes give rise to can easily make familiar narratives empathically inaccessible. Narratives according to which we are much like everyone else, or perhaps a little worse than average in some regards which we are sensitive about, fail to harmonise with the keen sense of our own importance and majesty. Making sense of this experience of one's own value and one's apparent past requires that magnetically repellent ideas ('my life has involved only very qualified successes and has amounted to very little', 'I am of great cosmic importance') are reconciled. One way for a narrative to achieve this is to revise the past. These self-narratives can become delusional when formed in response to an agent's need to preserve some reassuring picture of themselves against particularly challenging circumstances.⁵⁴ In the case of mania, this will often be a grand (or grandiose) picture of oneself. In depression, delusions of guilt can be motivated in the sense that they provide an intelligible explanation for why one feels so guilty.

Revisionary narratives can directly undermine agents' autonomy, and interfere with self-trust. When we rely on revisionary narratives, we rely on the world being a way other than the way it is. Depending on how far ranging our revisions are, we can find that our plans consistently end in frustration because we have not adequately appraised our abilities or context. This frustration means that we are less effective as agents, and so the agency on which our autonomy is founded is weakened. Our self-trust is likely to

⁵⁴ McKay et al (2005) suggest that delusional beliefs emerge as procedurally acceptable hypotheses which explain our abnormal experiences of value, but which are motivated by our refusal to accept that our experiences of value are unreliable.

be eroded as we encounter this evidence that we are ineffective in executing our plans or bringing our judgements to bear on the world (injuring type 2 self-trust). In the case of delusional narratives, however, our self-trust might initially be defended by our revisionary narratives. If we are excellent specimens, and very likely to succeed, our failures will require explanation, and so hypotheses which explain these recurrent but unlikely frustrations are adopted. However, as the delusion elaborates, we are narrating more of our past and present in revisionary ways, and so moving further and further from a reliable picture of the world. We are likely then to encounter more frustration, rather than less, (Bortolotti 2020, 70-98). Such an unreliable autobiography leaves us working on projects that are simply unrooted in the practical world that we are in fact acting on. Our agency becomes ineffective and our autonomy suffers. Should our delusions come to an end, our autonomy is likely to suffer further from a crisis of self-trust. Bortolotti et al (2012, 117) note, "[t]he effects of making one's self-narrative correspond to reality can be devastating, and many people experience severe depression when they recover or acquire insight into their illness." Even without these depressions, a sudden loss of self-narrative leaves one with insecure self-trust, and the rates of frustration or passivity that one's agential project involves might prompt (and warrant) active self-distrust.

Thus, we have an additional desideratum. Narratives cannot be unreliable or significantly revisionary. The self-trust that these narratives establish must be at least somewhat warranted. There is, however, flexibility here. As we saw above, self-narratives that are a little rose-tinted are valuable for establishing aspirational pictures of the self, which scaffold more effective and virtuous agency than painstakingly accurate narratives do. Narratives that are somewhat revisionary can be useful in keeping us engaged and in holding us to standards we set. We have three desiderata:

- 1) That our self-narratives keep us receptive to our own experiences of value as moral testimony,
- 2) That they help us secure (establish and maintain) self-trust, and
- 3) That the self-trust they secure is warranted.

In the next section, we will consider some narrative structures that have been suggested as advantageous to bipolar agents, and extract the most valuable structural elements.

§4.5 - Bipolar Narratives

Two sorts of narrative that have been offered as suitable candidates to make sense of a bipolar existential reality are *dangerous gifts narratives*, and *hinge narratives*. Dangerous gifts narratives analogise the sensitivity and ranging experiences of value that BoRD involves to a valuable gift that can be dangerous if not correctly managed. Structural features of dangerous gifts narratives will vary depending on how the danger is imagined, and in what sense it is a gift. A prominent iteration of this narrative emerged from the work of activists involved in the Icarus Project, (McNamara & DuBrul 2007), who take the artificial wings Icarus flew on to escape Crete as the metaphorical gift. Hinge narratives are suggested by Nancy Potter (2013), as structures that might be offered by therapists helping bipolar patients narrate their experiences in more fruitful ways. Both of these narratives involve structural features that go some way towards meeting our desiderata, but neither satisfies all three. We shall examine these narratives and carry forward what is most useful into our discussion of *medicine journey* narratives.

§4.5.1 - Dangerous Gifts

The Icarus Project brought together people living with various sorts of madness, including many activists, to cultivate better understandings of madness and illness. The narrative that this project trades in analogises madness to Icarus' short-lived gift of flight. Icarus' father gave him wings made of feathers and wax so that they could fly to freedom from their prison on Crete. If they flew too low to the sea, the spray would soak the feathers and the wings would be too heavy to fly. If they flew too high, however, the heat of the sun would melt the wax and the wings would disintegrate. If they did not keep to the safe medium space, they would crash into the ocean. Founder of the project, Jacks McNamara, draws the comparison as follows:

[S]ometimes the most incredible of gifts can also be the most dangerous. With our double-edged blessings we have the ability to fly to places of great vision and creativity, but like the mythical boy Icarus, we also have the potential to fly dangerously close to the sun—into realms of delusion and psychosis—and crash in a blaze of fire and confusion. At our heights we may find ourselves capable of creating music, art, words, and inventions which touch people's souls and shape the course of history.

(McNamara & DuBrul, 2007, *i*)

Madness in general, and bipolar disorder in particular, are susceptible to this narrative. The manic sense of things is a dangerous gift insofar as it offers the kind of vantage and artistic insight that McNamara and DuBrul discuss, but also because if it is not carefully managed, it can turn psychotic in ways that send us 'crashing into the sea'. We could also say the same for depression. Whilst it is less attractive than some manias, Solomon shares the reports of people who have experienced severe depressions and attribute their valuable moral and interpersonal sensitivity to that experience (2008, 528-9).

How does this narrative fare against our desiderata? Certainly, it keeps agents open to the manic sense of things. The value in that sense of things, explored in the previous chapter, is what makes this experience a *gift*. This narrative does not need to involve any revision the way that delusional narratives often do, and indeed warns against following manic lines of thought too far and too uncritically. The flight image instructs bipolar agents to continually keep themselves aloft, in the habitable zone, emphasising the need for continued effort to maintain one's position. Such sustained self-management makes it more likely that any self-trust we enjoy will be warranted, to at least some degree. (We address how self-management strategies protect and scaffold autonomous agency in the next chapter.) The dangerous gifts narrative, therefore, seems to satisfy desiderata 1 and 3.

Before considering desideratum 2, let us note the particular features of this narrative that cause it to meet our desiderata. The value that makes mania a (dangerous) gift is described in terms of vision, and the ability to act upon the world with the advantage of that vision. Artistic expression makes up the bulk of McNamara and DuBrul's examples, but what underpins this artistic ability is a perspective from which one can see the value of things. The narrative is one that draws the agent's attention to the moral knowledge that we established, in the previous chapter, is both valuable and which the manic agent has certain privileged access to. It is an additional advantage of this narrative that it satisfies the first desideratum by focusing attention on moral knowledge, it fills the variable 'gift' with a well-chosen candidate.⁵⁵ This narrative is not self-deluding. We will preserve this when discussing medicine journey narratives. The third desideratum is satisfied by the narrative's emphasis on self-management and avoiding the extremes in which we are more at risk of revising our narratives in delusional ways. This ensures agency warrants trust, rather than specifically preventing revision. Indeed, as long as this self-

⁵⁵ Certainly, most people who trade in dangerous gift self-narratives are unlikely to think in terms of moral knowledge, but priority given to "a different kind of insight, compassion, and expression of the human condition" (Jamieson 1994, 102) focuses attention on the right things, even if under another description.

management is conducted effectively, a degree of revision of the sort Bortolotti and Summers discuss could be accommodated without self-trust becoming unwarranted.

What is not clear is that this narrative secures self-trust. Icarus' wings are the gift, and their fragility is what makes them dangerous gifts. But the structural feature that capacitates the agent, keeping them engaged in their agential project and in the ongoing practice of self-management, is the same feature that endangers self-trust. The agent might, at any time, find themselves outside of the safe zone. Winds or thermals might act on the agent as they fight to remain neither too high nor too low, which suggests fragility in this self-management. At any time, a stressor or an unpredictable affective shift might be severe enough to force one out of the habitable zone. In this situation, the wings melt or soak, and the agent crashes into the sea. There is an all-or-nothing element to narratives, according to which agents must constantly strive to keep themselves from disaster, and where disaster is conceived of as terminal. Self-narratives that cast failure as catastrophic present a high risk that we will suffer catastrophic loss of self-trust when we fail by the lights of these narratives. It is misleading to tell bipolar agents that if they fail in careful self-management or suffer a serious episode of psychosis, that they necessarily crash and burn. We can, and do, survive and recover from these episodes, but do so more easily when our self-narratives emphasise our capacity to pick ourselves up and dust ourselves off. Dangerous gifts narratives, at least the variants that cast the dangers as catastrophic and terminal, render our self-trust *fragile* in undesirable ways.⁵⁶ They can establish self-trust, capacitating agents by ascribing gifts, to them, but if the dangers of these gifts are dangers we cannot recover from, then they make it much harder to maintain self-trust.

§4.5.2 - Hinge Narratives

Potter offers hinge narratives as meaning-making frameworks that might allow bipolar agents to make sense of their experiences where dominant sorts of illness narrative (restitution narratives and quest narratives, per Frank 2013, 115) are unsuitable due to the instability discussed in §4.3. These narratives

⁵⁶ Taleb understands fragility as the property in virtue of which systems fail to survive 'the unexpected'. It is the undesirable extreme of a spectrum that runs through desirable robustness (being unaffected by the unexpected, to yet more desirable *anti-fragility*. Anti-fragile systems become more powerful, effective, or secure when they encounter the unexpected, (2012). Good self-narratives are at least moderately anti-fragile, securing deeper self-trust the more often the agent successfully handles a situation they had not expected or prepared for. This is valuable only to the extent that deepening self-trust does not create complacency. Evidence that I can handle the unexpected should not cause me to trust that I don't need to plan for the unexpected. It is for this reason that I do not include anti-fragility specifically as a desideratum.

describe bipolar agents as being on a hinge that can swing to be more or less open to the creativity and free play of mania. Potter deploys the image of a stable door that swings in two directions, and is open on the top. These features are included in order to secure certain desirable features of the narrative. A hinge allows the agent to understand themselves as more or less open, at different times, to their experiences of value. The open top of the door ensures that the agent is never wholly closed off. Importantly, Potter highlights that this door can be moved by the 'winds' of one's mood, but also by other people who push the door in either direction. In order to make sense of their experience, bipolar agents can consider how open to manic or depressive experiences of value they are at present, and whether they need to adjust this. If they are so closed that they do not feel they can be creative, they should perhaps open the door somewhat, however if they are wide open and exposed, they should close themselves somewhat.

The organising metaphor of the hinge narrative allows us to talk about being open to the manic sense of things, and importantly, the degree to which we are open to that sense. Indeed, as Potter presents it, the narrative is likely to be useful for agents who are already open to their experiences of value and are concerned not to become closed to them. The narrative avoids the particular risk that dangerous gift narratives presented to our second desideratum, on which dangers were presented as calamitous and beyond recovery. The hinge narrative correctly suggests that we might, sometimes, be dangerously exposed to the manic or depressive sense of things but presents this as a situation that can be rectified by closing ourselves more to that perspective. There is a failure state, but not one that is final or terminal. We can move from the failure state back into the success state, becoming open to the manic sense of things to a safe degree. This satisfies the first desideratum, without creating the problem that made the dangerous gifts narrative unsatisfactory against the second desideratum (which we shall return to in a moment).

The hinge narrative offers more indirect support in satisfying the third desideratum. The narrative places particular emphasis on the role of others in our social environment in managing how open or closed we are to these kinds of normative outlook. When we narrate our pasts in revisionary ways, we often encounter rejection or non-recognition from others, especially when the narratives we offer are

revisionary.⁵⁷ This narrative carves out space for a check on the revisionary tendency that is particularly acute during manic or depressive episodes, by reminding us that our degree of openness to the manic or depressive sense of things is rightly answerable to peers in our social context. This attention to the standing our peers have to check our perspectives supports the attitude of self-answerability that we noted in Chapters 1 and 2.⁵⁸ Such attention to external perspectives is a useful check on the revisionary tendency, and so supports satisfaction of our third desideratum.

These are, however, qualified satisfactions of our desiderata, and worries loom. The hinge narrative assumes that the bipolar agent is already open to the manic sense of things, rather than specifically sustaining that openness. It does not include structural features that ameliorate stress or shocks, or anything else that causes an agent to reconsider their relationships with their own experiences. This does not demonstrate a weakness to this self-narrative, but rather highlights its localised utility. The narrative is offered for cases in which the first desideratum is met independently of the agent's self-narrative. Regarding the second desideratum, the hinge narrative supports self-trust by casting the dangers of overexposure as fixable rather than terminal. This secures self-trust in the sense of *maintaining* it, but not in the sense of *establishing* it. This self-narrative does not include features that capacitate the agent, the way that the dangerous gifts narrative does, for instance. Finally, the importance placed on social checks on our narratives should not be understated. The hinge narrative includes a check that is likely to support more effective agency, and to make what self-trust the agent enjoys more warranted.

The narrative is, however, somewhat limited in its intended scope. The agent can make sense of their experiences of value at this or that time in relation to their openness (or closedness) to the manic or depressive sense of things, but the organising metaphor of this narrative does not lend itself easily to

⁵⁷ This is an important, but imperfect check, relying on more than the memories of others. Our narratives might be rejected because others insist on remembering our deeds under act-descriptions that we do not recognise, and there is no guarantee that the act-descriptions our peers offer are more fitting. Indeed, insistence on particular act descriptions is a noted mechanism of oppression in unjust social relations (Bierria 2014, 2020, Radoilska 2021a). I address contested act-descriptions in Chapter 6.

⁵⁸ Recall, self-answerability is necessary for autonomy on Westlund's theory (2009) and amounts to a commitment to regulative guidance, ensuring that our emotions and evaluative judgements continue to count as reasons-responsive even when they get the responses wrong.

narrating more complex relations. We can vary significantly in how opened or closed we are to our own sense of things at different times, and in different contexts, and whilst it is valuable to know that we are, we need a richer narrative to place these variables in sense-making juxtapositions. Becoming more open to the manic sense of things, or more stable in one's degree of openness, are transformations that can have sense made of them in terms of the motion of the hinge, but changes such as falling in love, the development of class consciousness, or the nuancing of one's youthful idealism, are not. The hinge narrative is offered for use in therapeutic contexts when bipolar agents are in the process of recovering from crisis, or of building more flourishing lives. These are often contexts in which some progress has already been made, and ongoing support is present to take up some of the work our desiderata expect of a self-narrative. Hinge narratives are specialised sense-making tools suitable for their particular work, but cannot be successfully generalised beyond this context.

The hinge narrative has structural features that help us maintain self-trust, by presenting the threat to our appropriate openness as something we can recover from, rather than a fatal crash and burn. The narrative also creates a particular place for social interaction which restrain the tendency to spin revisionary narratives. We shall preserve these features when considering the medicine journey narrative. The hinge narrative lacks, however, features that establish self-trust, before maintaining it, and which keep us invested in the efficacy of our agency in ways that do not render our self-trust unwarranted.

§4.5.3 - All Told

Dangerous gift narratives allow us to preserve a positive self-image, according to which we are justified in having high self-esteem and believing that our contributions are worthwhile. Hinge narratives help us find in-the-moment self-knowledge regarding how much we are allowing our experiences of value to inform our thinking and action. Neither wholly satisfies our desiderata, but both include structural features that are valuable for satisfying our desiderata. We should remember that these narratives are not valuable only insofar as they contribute to satisfying these desiderata, of course. Many bipolar agents will find these narratives valuable, with respect to their autonomy or some other value, in virtue of precisely those structural features I have described as limitations. Different agents establish autonomies of different shapes, and whilst this chapter is aimed at solving a general problem of self-

trust, it is perfectly likely that other bipolar agents will solve that problem in other ways, and find quite different uses for the features I that now set to one side.

§4.6 - The Medicine Journey

In this section I argue that the *medicine journey* is a narrative structure that will satisfy our desiderata. The medicine journey is an early example of a monomyth, the structure proposed as the underlying skeleton of a great many culturally important stories, from folk tales to ancient mythology to the texts of living religions.⁵⁹ The hero of this narrative begins in a safe or familiar world, but due to some calamity or sickness (their own, that of a loved one or of the community) is compelled to travel across a threshold into an unfamiliar world of danger. In this world they encounter challenges and discover, or are awarded, an elixir. The hero then returns with the elixir, with which they restore the sick person or community.

The medicine journey is a quest narrative, echoing MacIntyre (1985, 219), of the sort noted in §4.2.2. The narrative is teleological: the bipolar agent is in the unfamiliar world in order to find the elixir, the elixir is the point of being there. The medicine journey has an underspecified goal. The elixir is under-characterised, only gaining specificity during the journey. Whilst I argue that there is esoteric moral knowledge to be found in the manic experience, no particular piece of esoteric moral knowledge is what makes the perspective valuable. Crucially, we find ourselves deposited in the midst of this quest: “characters of course never start literally *ab initio*; they plunge *in media res*, the beginnings of their story already made for them by what and who has gone before.” (215). The narrative projects meaning backwards even if we did not enact particular narrative steps such as choosing to leave familiarity or search for anything. Similarly, narratives project into the future, their teleology making particular futures narratively fitting in ways that can daunt or inspire us. The medicine journey narrative makes

⁵⁹ The medicine journey served as the template for the now more well-known *hero's journey* (Campbell 1968). The hero's journey is indeed a medicine journey with more stages specified and a wider range of variables suggested to be the impetus to depart or return. Prior to this, the medicine journey perhaps received its most noted examination by Eliade (1988).

sense of our experiences as a quest for moral knowledge that we, routinely, find ourselves already embarked on.⁶⁰

What structural features of this narrative are advantageous? The medicine journey as applied here casts the bipolar agent as a boundary-crosser, who moves between the familiar world in which their experiences of value are similar to those of their peers, and the unfamiliar world in which their experiences are esoteric and are not easily accommodated by received hermeneutic resources. In the unfamiliar world, they can find an elixir in the form of the esoteric moral knowledge that we discussed in the previous chapter (variously glossed as a unique sensitivity, artistic insight, creativity, etc) which is valuable for them, their loved ones, or their community. The experience of BoRD, then, is cast as a mission to find and bring home something that is valuable, but which one does not otherwise have access to in the familiar world. (That such an object of value that is otherwise inaccessible does in fact exist, as argued in Chapter 3, prevents this narrative structure from amounting to self-deception.) Bipolar agents might conceive of themselves as bringing the elixir back to their peers in the familiar world, but this isn't necessary and there are reasons to avoid adopting narratives according to which the bipolar agent ranges into the unfamiliar world *on behalf of* those in the familiar world. The unfamiliar worlds of mania and depression are dangerous and distressing, and no community has a claim on a bipolar agent that they undergo such travails, neither does any edification of the community *justify* the bipolar agent's suffering. All this narrative need to is make the experience meaningful and navigable.⁶¹

Whilst many iterations of this structure in folk tales and myths are single stories, the medicine journey as bipolar narrative is a cyclical story. As bipolar agents have recurring episodes, they will find that they are in different parts of this journey at different times. The agent who had one serious episode but has been in remission since can be prepared for their next one, knowing that they made it home last time. The agent whose episodes are far more frequent can understand themselves as the person who goes on these travels by habit, or by dint of their role. This is why it is important that we find ourselves *deposited*

⁶⁰ This is distinct from Frank's (2013) usage of the term 'quest narratives', as a subset of illness narratives. Frank's (illness) quest narratives conclude where we conquer illness as the bodhisattva does, surviving through our persistence (rather than the way Heracles does, by forcing opponents into submission, 119). This is one way that a bipolar agent might self-narrate, but the medicine journey is not confined to this sort of conclusion.

⁶¹ Whilst above I noted that the medicine journey is teleological, the teleological relation can be normative without being justifying or demanding. The narrative involves the more modest presentation of one's being in the unfamiliar world as *intelligible* with reference to a search for moral knowledge.

on this quest, already embarked without having to choose to leave. Whilst this would prevent us making a tight chronological sense of our experiences, that is not the kind of sense we need to make. It is sufficient that we can understand what we were doing at different times as ranging into the unfamiliar, as learning the thing that was valuable to learn, or returning with that knowledge to the familiar world. The narrative relations between these stages of the cycle do the sense-making work where causal or chronological relations would not.

This narrative incorporates the structural features we selected to preserve from previous narratives. From the dangerous gifts narrative, we preserve the focus on vision, or moral knowledge. The boundary-crosser can travel into the unfamiliar world and see things that their peers in the familiar world cannot. The elixir that they return with is the same object of value that makes madness count as a dangerous *gift*, (esoteric moral knowledge, per Chapter 3). From the hinge narrative, we preserve the role of peers in the agent's social context as a check on self-interpretation and on interpretation of the moral knowledge that is found. According to the medicine journey narrative, the point is to bring the elixir back to the familiar world, where our perspectives can be easily shared with others. Consequently, the bipolar agent must consider how they will offer what they have learned as moral testimony, to people who do not have access to the unfamiliar world. Reflecting on this question forces the bipolar agent to consider the perspectives of those they will interact with and so introduces the valuable check on interpretation that the hinge narrative secured.

Some particular features of the previous narratives were valuable particularly in conjunction. From the dangerous gifts narrative, we hoped to preserve the capacitation of the agent which established self-trust, and the attention paid to ongoing self-management needed to keep oneself in the safe zone. From the hinge narrative we hoped to preserve the non-calamitous view of failure, according to which we can get up and dust ourselves off, rather than being doomed should we ever find we have flown too close to the sun. The dangerous gift narrative kept us focused on self-management in ways that will likely make our self-trust warranted, but made what self-trust we do enjoy fragile by presenting failure as total. The hinge narrative helped us maintain self-trust by presenting failure as non-total, but did not give us an initial impetus to self-trust.

The medicine journey manages to unite these desirable features. It secures self-trust by casting the bipolar agent in a role: the boundary-crosser whose task it is to bring elixir home. Being cast in this role

capacitates the agent in the way we discussed in §4.2.2. Finding oneself on this particular quest, ranging into the unfamiliar world and returning, ascribes the relevant capacities to the agent to survive that place and to find their way back. The cyclical nature of the quest supports this capacitation. The bipolar agent who experiences successive episodes is someone who has made this journey before and can find their way back. The capacitating role of this narrative establishes initial self-trust, and the cyclical structure that means we are always (even if the long way round) returning home, supporting the maintenance of trust. The metaphor of a journey is more useful than that of flight, as if we ‘get lost’ we can at least keep going until we find the road again, as opposed to crashing into the ocean. The failure state suggested by this narrative (getting lost) is serious enough to focus the mind on the task of self-management, but is not terminal in the way the Icarus-styled dangerous gifts narrative was. Thus, self-trust is more effectively maintained. The medicine journey establishes, and continually re-establishes, self-trust without rendering it fragile, thus it more effectively secures self-trust across the immediate and longer terms.

So far, we have drawn comparisons between the medicine journey narrative and the previous, dangerous gifts and hinge, narratives. The medicine journey narrative preserves the structural features that were most desirable from the previous narratives and manages to balance both establishment and maintenance of self-trust. With these considerations in place, we can now evaluate how well the medicine journey narrative fares against our desiderata:

- 1) That our self-narratives keep us receptive to our own experiences as moral testimony,
- 2) That they help us secure (establish and maintain) self-trust, and
- 3) That the self-trust they secure is warranted

Against the first desideratum, the medicine journey fares well. It casts the agent as occupying a role in which it is their task to find the elixir and bring it home, be it for their own edification or also for that of their community. As I offer the narrative, this elixir is identified with the esoteric moral knowledge that the manic sense of things is particularly useful for accessing. Importantly, as it is a *search* for the elixir, this narrative focuses the agent on the task of distinguishing which parts of their experience are valuable moral knowledge and which are not. Disentangling reliable moral insight from the mistaken or illusionary picture of value that emotional experiences give rise to is not an easy task with relatively familiar neurotypical dispositions, and is not easier in mania or severe depression. The prospect of delivering the elixir to one’s community focuses the attention on the perspectives of one’s peers and so

prompts reflection on the interpretation of these experiences and how one would respond to scepticism regarding this moral knowledge. This ensures that agents are open to the manic sense of things, but not uncritically open.

The medicine journey satisfies the second desideratum by securing self-trust. The narrative casts the agent in a role defined by a task: traversing the dangerous bipolar world and returning home with specialised moral knowledge. Being cast in this role capacitates the agent as we saw above, and as the narrative is cyclical, continually re-capacitates the agent. Bipolar agents who deploy a medicine journey narrate themselves as the person whose past and future involve successfully navigating the unfamiliar world, and so create the expectation that they will successfully do so this time. In capacitating the agent, this narrative secures trust in the procedural soundness of their capacities. Due to its cyclical nature, it suggests that the agent has brought something valuable back from the unfamiliar world during their last episode (even if they have not yet clarified to their own satisfaction precisely what it was), and so establishes trust in the substantive appropriateness of their sense of value. These are types 1 and 3 self-trust as described by McLeod and Sherwin. Securing type-2 self-trust is trickier, as BoRD does create unquestionable obstacles to agents executing their plans and acting on judgements made before and after affective shifts. However, as all narratives project into the future and so create expectations that can found self-trust, the medicine journey's cyclical nature allows us to trust more in our own reliability. The narrative involves a homecoming to the familiar world where it is easier to rely on our agency, and where we can update our existing plans and commitments with the new moral knowledge we have found. Whilst we might feel we are lost in the unfamiliar world at some particular time, the cycle is not complete until we return home, and so that safer environment, where we can be more reliable, is always placed on the horizon. This projection into the future gives us an aspirational, even if epistemically irrational, focus upon which to establish type-2 self-trust.

The third desideratum is that self-trust be warranted. By placing prominence on the return home to the familiar world, where perspectives are closer together, the medicine journey introduces a social check on agents' interpretations. That the agent understands themselves as returning to a place where effective communication is expected finds their attention is drawn to how their new moral knowledge will be received. This check, drawing on other perspectives, creates opportunities to catch procedural errors, and so agents are less likely to make them (warranting type 1 self-trust). Genuinely considering alternate views on value better informs our own perspectives and allows us to synthesise our moral

outlook. We stand a better chance of making substantively appropriate judgements if we are exposed to, and sensitive to, other perspectives, and so the substantive appropriateness of our motivations and interpretations are safer, (warranting type 3 self-trust). By offering new moral knowledge as moral testimony, we are likely to make the same implicit commitments Summers noted (2017) which create expectations in others and ourselves that we will continue to express these values in our behaviour and future rationalisations. Thus, the pressure to be consistent makes us more likely to live up to the standards we set for ourselves and so we can increasingly come to rely on ourselves living out these values (warranting type 3 self-trust). We are, therefore warranted in the three kinds of self-trust that McLeod and Sherwin outlined.

§4.7 - Securing Autonomy by Self-Narration

BoRD creates a problem for autonomy by undermining the self-trust agents require to enjoy tolerable degrees of autonomy. The episodic nature of BoRD creates, for agents, stark self-disagreements on matters of value. The agent must be wrong at least once in these disagreements, but from the inside has no way of determining when. Agents might attribute the mistake to their affective episode, taking a pathologizing stance which disregards the manic sense of things as moral testimony, but will find themselves confronted with a general undercutter to their moral knowledge. Consequently, their trust in their own capacities to track and respond to value (types 1 and 3 self-trust) is arrested. Self-narration, on the template of the medicine journey, allows agents to secure self-trust in McLeod and Sherwin's three regards.

This self-trust supports agents' autonomy in the ways we explored in §4.2.1. We shall rehearse these here as we examine the consequences such trust-supporting narratives have for the shape of bipolar agents' autonomy. The medicine journey capacitates agents in a way that creates expectations that they will navigate an agential project through the various mood states they will experience. Creating this expectation makes it more likely that bipolar agents will exercise the capacities that their attention is being drawn towards. Our agential capacities are the foundations upon which our autonomy is built, and so by exercising these capacities, we extend the foundations of our autonomy. The degree to which this makes our overall autonomy more robust depends on how circumstances allow us to build on that expanded agential footprint. All things being equal, we enjoy wider autonomy when we build on an

expanded foundation, (compare with James' account (Weiner 2011, 471) in §3.2.3). These capacities are also made more effective by the scaffolding achieved by a robust self-narrative. By creating expectations that we will succeed, this self-narrative keeps us invested in our agential project and therefore likely to make progress on at least some fronts. The priority given to communicating mad perspectives to peers in the familiar world ensures that we interact in our social environment. This introduces valuable checks, as our peers troubleshoot our plans and offer valuable perspectives on our interpretations that will help us get a firmer grasp on the values we are considering. Sharing esoteric moral knowledge creates the expectations in others that we will express the values we describe, and these expectations scaffold effective agency for us. Trusting in our own perspective makes us more effective in the exercise of these agential capacities, providing stronger, as well as wider, foundations for our autonomy.

The most important sense, however, in which the medicine journey supports autonomy, is in leading us to exercise autonomy, so that it can extend further in various dimensions over the expanded footprint noted above. When we trust that our own motivational systems are reliable, in the sense that they do not tend to make procedural errors, and in the sense that they tend to produce appropriate conclusions, we will have the confidence to close deliberation and commit to some project. Making commitments extends our autonomy further in certain dimensions. The more commitments we make and fulfil, the further our autonomy extends in the dimension of self-determination.⁶² This is an exercise of agential capacities but is also an exercise of autonomy. We shape our life, and the kind of person we are, by committing to certain projects or to being certain ways. Self-trust is crucial if we are to move from the exercise of the agential capacities we use in deliberating and appreciating, to exercising the autonomy-constituting capacity to self-determine. Agents who lack this trust see their autonomy injured. Either it atrophies, extending only a short distance in the self-determination dimension, or it is defeated by passivity. In the former case, our practical life becomes small, and for want of commitments that make us into a particular kind of person, we fail to live autonomously to the extent that ours fails to be a life where autonomy could belong. In the latter case, we will live a practical life and be some kind of person, but specifically we will be a compliant or acquiescent person. We will 'do the done thing', or find that we are relying on other people to direct our agential project for us. Whilst this does not preclude us having a rich and complex agential project that is satisfying, it is not an autonomous life we are living,

⁶² Self-determination is the example dimension we shall explore next chapter, chosen for the sake of having an example with which to examine the shapes of autonomy. (See §0.3.4)

as we are relatively passive in determining what we do. Indeed, this is the private version of the key failing on social models of autonomy, where we fail to be autonomous because we are bound, enslaved, deprived of opportunities in our social context, or kept ignorant of the range of possibilities we might pursue. The difference is that it is us who enforces these restrictions, out of fear, a lack of self-respect, or as we have examined, a lack of trust in our own motivational systems.

The narrative's teleological objective, whether we gloss it as moral knowledge, great vision, or artistic insight, supports the foundation of our autonomy in relational senses. The search for moral knowledge keeps the bipolar agent invested in considering what reasons are apparent to them, whether they are real reasons or merely putative reasons, and how best to interpret the values that underpin these reasons. This attention to their own view of the normative landscape ensures that the agent continues to take an agential stance towards themselves, and so offers protection against the pathologizing stance on which their experiences are explained rather than read as testimony. Regarding one's experiences as testimony means that the agent's self-relation can be a site of the recognition that is a basic component of autonomy on views that take autonomy to be constitutively social. This stance towards one's own agency was the crucial first desideratum for the medicine journey self-narrative because satisfying it is necessary to meet the relational conditions for autonomy on constitutively social views. Taking this stance towards ourselves also means we are more likely to present our view as testimony, in situations that call for moral testimony, and under descriptions that make clear that it is moral testimony. In so doing, we prompt our peers to recognise our agential capacities at work, and so extend social recognition to us.⁶³

Self-trust is a crucial element in the construction of a robust autonomy. It creates the space in which we exercise our agential capacities and in which these capacities are more likely to be effective, laying the practical foundations for our autonomy. It prompts us to take an agential stance towards our selves, ensuring that we recognise those capacities for what they are and recognise them at work when they are, which in turn causes us to prompt others to take such a stance towards us, satisfying the relational conditions of autonomy. But beyond the foundations, self-trust is what smooths the transition from it seeming to us that some reasons are our best, and our committing to do something or be some way.

⁶³ Recognition in the self-relation is, indeed, necessary to meet the relational conditions of autonomy. External social recognition is needed, but cannot get us very far in establishing robust autonomy if we do not extend that recognition to ourselves.

Self-trust is the currency with which we exercise our self-determination. This self-trust is secured by a reliable self-narrative that establishes and maintains our self-trust. In the face of the challenges to self-trust presented by BoRD, bipolar agents can rely on the medicine journey narrative to make sense of their shifting experiences of value in a way that capacitates them, establishes self-trust, and maintains it through these transitions.

§4.8 - Conclusion

The medicine journey is a fairly detailed central metaphor that allows us to speak about a wide range of self-narratives which bipolar agents might adopt. The structural features I have highlighted as valuable here include a teleology that capacitates the agent, cyclicity which puts 'home' (familiarity, less distress, success) always on the horizon, and focuses attention on succeeding at a task without presenting failure as terminal. The medicine journey is an 'off-the-shelf' narrative structure that includes these features in particularly useful ways, but what matters for autonomy is not that bipolar agents adopt medicine journey narratives, rather it is that bipolar agents' practices of self-narration deploy structural features in ways that support their autonomy. The medicine journey is a valuable resource for this, but it is not necessary that the structural features are arranged in this manner, or even that these are the structural features used to secure self-trust. The medicine journey is an attractive candidate, but not dictum.

I have argued in this chapter that self-narratives that include these structural features are particularly more likely to secure bipolar agents' self-trust whilst also ensuring that they continue to take their own experiences of value seriously as moral testimony, and avoid the kinds of revisionary narrative that would make their self-trust ultimately unwarranted. Self-trust facilitates the establishment of robust personal autonomy, and is in specific regards constitutive of that autonomy (involving, as it does, the agent's recognition of their own agential capacities). A particular dimension in which autonomy is likely to extend further as a result of robust self-trust is the dimension of self-determination. Agents who trust themselves are more likely to perform self-determining exercises of autonomy. Agents with high degrees of self-trust are not only likely to enjoy more robust or extensive autonomy, but are perhaps more likely to enjoy autonomy of particular kinds of shape, as the self-determination dimension will extend further than it does for people with lower degrees of self-trust. This is not to say that autonomy

in other dimensions is not also made more extensive by self-trust. As noted in §0.3.4, for the purposes of this enquiry I do not adopt a particular theory of autonomy with a view of which dimensions make up autonomy. Self-trust will tend to produce a variety of shapes on different theories. What is indicated here is that the topology is a key variable in autonomy between people, and that there is a plurality of shapes autonomy can take. The oddly lumpy autonomies of bipolar agents count as autonomy, even if they are harder to recognise prior to detailed discussions of where bipolar agents' agency lies and is exercised, and at what resolution it is shaping their lives.

Self-determination is a dimension of autonomy in which the efficacy of our commitments is crucial. Whilst self-trust is an important enabling condition for the extension of our autonomy in this dimension, that extension also requires a degree of success. We do not self-determine when we flip-flop between contrary commitments, or when our commitments come to mean very little for what we do. If our agential project is not in fact shaped by the commitments we make, we do not enjoy high degrees of self-determination. This is a particular problem for bipolar agents, again due to the episodic nature of BoRD. The next chapter will examine how this problem threatens the efficacy of bipolar agency and the degree to which bipolar agents can self-determine.

Chapter 5: Living Between Normative Worlds: Predicating Self-Determination at Scope

This chapter has been published in Philosophy, Psychiatry, & Psychology, 29(3):133-145.

§5.1 Introduction

Bipolar or related disorders (BoRD) present unique practical and existential problems for people who live with it. All agents experience changes in the things they care about over time, however people living with BoRD face drastic shifts in what seems valuable to them which frequently upset the plans and commitments that they previously determined they would pursue. Navigating these evaluative high seas presents agents living with BoRD with a distinctive existential problem which is not shared by those on calmer waters. This problem is how to sustain an agential project between and across the episodes of mania and depressions which restructure agency, shift normative outlooks, and change what the agent considers to be profoundly valuable. In this chapter, I draw out two contrasting ways in which an agent might scaffold their agency in the face of this instability in order to secure autonomy in its self-determination dimension, and manage the existential problem they face. I focus on particularly on self-determination because it, or something like it, appears as a feature or dimension of autonomy on a wide range of theories and it captures features that are widely considered to be central to personal autonomy, (making choices, non-passivity, authoring one's own life or one's own 'self' in relevant senses).⁶⁴

Someone living with BoRD might seek to support their self-determination in the face of BoRD by crafting appropriate self-management regiments. The first strategy involves managing one's affective episodes so that they do not interrupt one's plans and long-term agency over time. The second allows one the greatest degree of freedom in responding to changes in one's experiences of value. What distinguishes these sorts of self-determination is the scope at which they are predicated. Thus, I describe the former strategy as securing self-determination predicated widely, (or wide-scope self-determination), and the latter as securing self-determination predicated narrowly (narrow-scope self-determination). The

⁶⁴ See, for example, Colburn (2010, 22-31, as 'commitment'), Radoilska (2012b), MacKenzie (2014), Killmister (2017, as 'self-definition'), Pugh (2020, 8f39).

difference is in the scope of the determination. Commitments which guide us in more future decisions, requiring more steps to be fulfilled, and hedging out more possible options, are wider in their scope than those which range over smaller portions of agential life.⁶⁵ Whilst a determination can be wider in scope than some, but narrower in scope than others, I will speak of wide and narrow scope self-determinations throughout, for the purposes of clarifying the significance of this distinction. Scope, here, does admit of gradation, but this will not be significant for our discussion. Whilst both wide and narrow scope self-determinations allow an individual to rule themselves, they alter the overall shape of one's autonomy in quite different ways.

In Chapters 1 and 2 we discussed how agency is altered or wholly restructured by the affective shifts that characterise BoRD. In the previous chapter, we considered how agents might support their autonomy in light of the episodic nature of BoRD, particularly with respect to their self-trust. The impetus to exercise agency or autonomy depends on having a suitable degree of trust in oneself, and without this self-trust, agency lies fallow and autonomy fails to extend. Our purpose in this chapter is to examine how bipolar agents might make their agency successful, and their exercises of autonomy effective, in the face of persistent affective shifts. §5.2 will outline the practical and existential problem that affective shifts pose for bipolar agents' autonomy. This problem is that even if one has sufficient self-trust to exercise autonomy, one cannot rely on their future selves having a suitably similar normative outlook to carry out and fulfil the commitments one makes when exercising their autonomy. §5.3 will outline self-management as a method of scaffolding agency that is commonly practiced by neuro-atypical agents, and will outline two broad approaches to self-management bipolar agents might adopt for the sake of managing this existential problem. These self-management regimes will be organised around protecting self-determination predicated either widely, or narrowly. §§5.4 and 5.5 go into detail on what self-determination involves at these different scopes, highlighting what is valuable about each, and what the costs of each are. Neither sort of self-determination can be secured without some drawbacks. Protecting self-determination at once scope frequently weakens or removes it at another. These sections together demonstrate that exercising and supporting self-determination

⁶⁵ We could think of these as determinations to do bigger or smaller actions, if we understand actions in the terms Steward offers, (2012). If we take actions as mereological sums of activity, and adopt a mereological universalist position, we can describe 'bigger' actions as sums of many 'smaller' actions, or of unindividuated activity. Wider scope self-determinations, which make determinations over larger portions of agential life, include more (actual or possible) activity in their scope than to narrower scope determinations.

predicated at either scope will have different sorts of consequences for the overall shape of an agent's autonomy, pulling at the fabric in other dimensions. In §5.6 we will consider the choices that bipolar agents face with respect to self-determination. The choice to protect self-determination predicated at some scope has far reaching consequences for the shape and texture of an agent's autonomy. Understanding this distinction, therefore, is necessary if agents are to make informed decisions about the shape of their autonomy. However, this analysis also emphasises the wide possibility space in which agents, bipolar or not, can make choices about the shape and texture of their autonomy. This latitude permits us to craft agential projects and kinds of personal autonomy bespoke to our own situations and existential realities. The diversity in kinds of autonomy which the existence of a particularly lumpy bipolar autonomy evidences, extends through the many ways in which these lumpy sorts of autonomy can be fine-tuned to suit different lives.

§5.2 Affective Shifts and Self-Determination

BoRD is characterized by manic, hypomanic, or depressive affective episodes during which an individual experiences elevated or extreme emotional states. These episodes typically remit and re-emerge many times through a bipolar agent's life. However, symptoms of depression or mania experienced outwith episodes are also common. Many bipolar agents adopt self-management practices that involve constant vigilance regarding their emotional state, in order to head off an episode whenever one is emerging. As we have seen, the evaluations an agent makes are altered by these episodes, with depression bringing despair or a sense of general futility, and with mania bringing intense excitement, cheer, anger, or urgency. These changes in evaluation can drastically alter the reasons one takes oneself to have, producing swerves in one's plans, but they can also be so pervasive that they deprive an agent of any stable point of view that is aloof from their disorder.

Moore et al (1994) relate the biography of Mr. M during a time when he took a succession of medication holidays, each one occasioning significant hypomanic symptoms. Two separately stable value sets emerged during this period. While medicated, Mr. M valued his family life very highly, whereas when he was unmedicated, he found his family dull and stifling and valued his artistic endeavours and his new

relationship much more (167). At the other extreme, Weiner (2011) highlights cases in which agents are deprived of any point of view which can be isolated from the symptoms of BoRD. James, a participant at a bipolar peer support group, remarked that “bipolar alone is not sufficient to explain every behaviour that you do—but neither is it not the explanation for everything that you do” (470). His symptoms touched his every experience of value and prevented any values from emerging that are not at least partly imputable to BoRD. None of his evaluations were above suspicion of being ‘the disorder’ at work. Reifying something like BoRD creates some other entity to which one might attribute one’s thoughts or evaluations, and so ambiguity over the self emerges, (Dings & Glass 2020, 339). James understands his bipolar disorder to lay claim to at least some of his experiences, yet these experiences feel no less his own.

This poses a particular existential and practical problem for agents living with BoRD. Before they shape themselves into any particular person, they must contend with the fact that they will, probably repeatedly, experience affective shifts that offer very different views of the normative landscape. Most agents determine what they will value and do so with the knowledge that, over time, they will acquire new information, endure transformative experiences, grow tired of old interests, and excited about new ones. Agents living with BoRD, however, must acknowledge that whatever they determine to value or do now, their experience of value will sometimes be radically at odds with their current one, that this will occur with little warning, and might remit as suddenly. This is an existential problem because becoming a certain kind of person takes longer than one’s view of the reasons to be that kind of person are likely to last. It is a practical problem because an agential project is a difficult thing to craft if one’s view of reasons to act will not be similar in key ways across time.

This problem concerns our autonomy as it is a problem of self-determination. Self-determination is a way of exercising autonomy on almost any theory, and is a distinct dimension of autonomy on Mackenzie’s relational account. She tells us that self-determination involves “the freedom and *opportunities to make and enact* choices of practical import to one’s life, that is, choices about what to value, who to be, and what to do.” (2014, 17, my italics). Our motives when we self-determine must be properly our own, even if we do not feel a phenomenal ‘mine-ness’ attaching to them. One way for our motivating values to be our own is if we determine for ourselves what our values are. This is an exercise

of autonomy, in that we use our agential capacities to take on these values rather than passively accepting them, or keeping them out of habit. Autonomy requires that one determines for oneself what to do, rather than acquiesces to “do the done thing”. How bipolar agents might self-determine is a difficult question as the normative outlooks they occupy at different times are phenomenologically of a kind; none identifying themselves as unauthentic or lacking authority. Often these episodes have precisely the phenomenology of a revealing or epistemically transformative experience, “the feeling of successful insight.” as Lodge puts it (2020). The problem is rooted in the phenomenal experience of value. The evaluations that one makes during an affective episode frequently do not come with any distinctive feeling of alienness. If indeed a phenomenal experience of ‘mine-ness’ is available to agents (prior to the kind of reflection Dings & Glass discuss, 2002, 335), it is present in these episodic evaluations.

On a wide range of theories of autonomy, some attitudes about value are taken as speaking for the agent, and as standing on the ‘self’ side of the self-determining relation. The self that determines is identified with, composed of, or spoken for by the privileged attitudes. An agent’s decisions, commitments, and actions are (generally) autonomous or (particularly) self-determining as long as they stand in some proper relation to the privileged attitudes. Deep or authentic selves are variously understood to be composed by a set of cares which enjoy higher order endorsement (Dworkin 1988, Frankfurt 1998, Bratman 2003, van Willigenburg & Delaere 2005, Shoemaker 2015, 44), as unreflective cares which thread the narrative of our lives together, (Jaworska 2007), or as clusters of different kinds of attitude which nonetheless combine to make up an agent’s deep self (Shoemaker 2015). Watson (1975) places agents’ reflective judgements about value in the privileged position, (see also, Colburn 2010, 25, Hieronymi 2013, and again, Shoemaker 2015). More recently, Killmister (2017) offers sets of *agential commitments* to populate this side of the relation. As noted in §0.3.4, the purpose of this enquiry is not to adjudicate between particular theories of autonomy, but rather to demonstrate the need to recognise bipolar and other sorts of neuro-atypical autonomy where they exist, equipping scholars to develop theories that accommodate neuro-atypical agents. For this reason, I remain broadly, though not entirely, theory-neutral between accounts of autonomy. I will refer to whatever attitudes that are taken to make commitments self-determining as *values*.⁶⁶ Several, although not all, of

⁶⁶ This is a suitably promiscuous term. Our values are evaluative, but beyond that, they might be any kind of attitude (we judge that our values are correct or fitting, we care about our values in both Frankfurt’s and

the theories noted above require that values be stable and reliable in the medium-to-long term if they are to authorise commitments as self-determinations. These are *longitudinal values*. I distinguish longitudinal from *episodic* or *contemporaneous values*. The problem that we address in this chapter is that bipolar agents cannot rely on, and might not possess any, longitudinal values. This is why the practical and existential problem noted above matters for an enquiry into autonomy. If longitudinal values are required for an agent to self-determine, then bipolar agents who lack such values cannot extend their autonomy in that dimension at all.

Agents with BoRD sometimes lack a stable value set which would, on various theories of autonomy, provide an authoritative self. Moreover, no experiences of value are phenomenally identifiable as better candidates than others for composing such a value set. Yet we must still take the normative outlook such agents have as speaking for them, at least during the period for which that outlook lasts. Whilst Jaworska's (2008) account of acting autonomously against our (longitudinal) values assumes *some* stability in the cares which compose a point of view, she allows that short terms of *caring* can do this work. Of a Mr. Lazaroff, in the last months of his life, she suggests that "a full gamut of [his] emotions has crystallized around the goal of staying alive" (101), and that this attitude speaks for him at that time.⁶⁷ Whilst the evaluations we make during these affective episodes are not values in the sense of being long-lasting, stable, and embedded in our day-to-day practices, they can compose for an agent "a stance she must acknowledge as her own" (being, as they are, "deeply rooted in the [agent's] psychic make up," (103). Such shorter-lived cares are episodic values. Kay Redfield Jamison puts the existential point well when she wonders which of her feelings, her crazed mania or her doomed depressions, are "her". She concludes: "[p]robably a bit of both, hopefully much that is neither." (1995, 93).

The values one determines for oneself must be those that speak for the agent, rather than ones which are imposed, and can be episodic values or Jaworska's cares, as easily as longitudinal values. One's

Jaworska's senses) or other suitable object, (good physicians take on, as agential commitments, the values that are important in their values-based practice).

⁶⁷ See also, Jaworska (1993). In both of these papers Jaworska considers fresh cares arising at the end of life, whereas in BoRD, fresh or episodic cares are likely to remit. A number of interesting implications of this distinction are worth exploring but require more attention than can be afforded here.

commitments must stand in a proper relation to those values, even if this requires that one revise some of one's commitments. Steps taken to ensure that our commitments are effective might injure self-determination when they curtail *opportunities* to act on new values (which might turn out to be episodic). Self-determination involves both *making* and *enacting* these choices, and BoRD sets these expressions of self-determination at loggerheads. Yet to conclude that a person with BoRD is without values, and therefore not a candidate for self-determination, would get us nowhere. We are investigating the kind of autonomy that is available to people living with BoRD, and what its ideal form is. Autonomy is still valuable and agents living with BoRD still face the difficult task of securing for themselves what kind of autonomy is available, and making decisions as to what that autonomy should look like. Where previously discussions of autonomy might have concluded that BoRD places a ceiling on how autonomous a life might be, recent multidimensional analyses invite more granular stories. Rather than to determine who is more or less autonomous, Killmister sought to measure the *shape* of a person's autonomy, (2017, 17–18). One's autonomy extends in various dimensions and might extend further in one than in another. This topographical exercise provides a much richer account and invites us to actively make choices about the shape of our autonomy. All agents confront the question of what shape to craft their autonomy into, and given the existential problem, agents living with BoRD must understand how BoRD confines the possible shapes of their autonomy, and which self-management strategies produce which shapes. As I shall argue in this chapter, there are also textural differences in autonomy that will be produced by agents' autonomy when cultivated in the face of different existential realities, will be the consequence of various shaping choices, and which might be the object of an agent's choice as they carve out their own autonomy.

Self-determination is one of a number of dimensions of autonomy, and understanding the various ways one's autonomy might extend through this dimension can helpfully reveal something about the shape and texture of manically depressed agents' autonomy. In order to recognise the excellent form of autonomy available to people living with BoRD, we must understand the shapes that autonomy might take. In the next section I will sketch two sorts of self-management and support regime that a person living with BoRD might adopt to support their self-determination. These regiments are configurations of commonplace self-management strategies, brought together to highlight the difference in the kind of self-determination they support.

§5.3 Two Self-Management Regiments

Self-management and support regiments are collections of strategies that people living with chronic illnesses and mental disorders adopt in order to manage the effects of their illness. A wide range of self-management strategies are practiced by people living with bipolar disorder (Depp et al, 2009), and many of them involve significant support from friends, family, colleagues, and clinicians, (Brahim 2018). Common self-management strategies for BoRD include identifying the triggers and early warning signs of an episode, avoiding triggers, attempting to manage oncoming episodes so that they do not become more severe, and maintaining access to the resources and support necessary for a swift recovery. Support from friends, family, and clinicians might include an understanding that a bipolar agent will sometimes be less reliable than at other times, or that they might withdraw socially and not do the work of keeping up friendships. At other times, it might involve friends imposing their company on someone whose self-isolation risks causing them to spiral. What support looks like will vary and depend on the needs, dispositions, and situation of each person. I will sketch two broad approaches to self-management one might adopt for the purposes of securing self-determination. Whilst these approaches are defined by the kind of self-determination they seek to secure, individuals who adopt either would still differ significantly in how they fill the variables.⁶⁸

§5.3.1 Self-management for Wide-Scope Self-Determination

This regime secures self-determination by making sure that when one makes a commitment to some project or end, one sees it through. BoRD creates obstacles to effective agency given that one will, eventually, have experiences of value which make the reasons underpinning one's existing commitments seem insufficient. One might become depressed, and no longer see the value in one's projects; or become manic, perceiving powerful new reasons that justify putting commitments on hold until one has written one's great novel. Consequently, commitments are revised when episodes or

⁶⁸ Self-management regimes are not typically constructed around one end such as self-determination, but around the range of valuable ends an individual has. These examples are simplified and exaggerated in order to provide an illuminating contrast. For a useful set of discussions around various facets of self-management, see special issue of *Philosophy, Psychiatry, & Psychology*, 27 (4), 2020.

symptoms alter one's normative outlook. Projects might lie fallow, or deteriorate due to neglect. One might expend resources, time, money, energy, on other projects, making it harder to do what one had previously intended to do. If one's commitments required particular things to be done at particular times, one might neglect or miss those opportunities due to no longer caring, or being more invested in some new manic fixation. This all makes it less likely that an agent's commitments come to fruition, and so threatens their self-determination. Agents are less likely to enact their determinations in the way relevant to autonomy.

Consider, then, what self-management might look like if it is geared towards overcoming this obstacle. Fairly rigid timetabling and diary-keeping, blocking off time specifically to let out whimsical or fantastic energy, and refusing impromptu invitations to concerts, dinner, or the pub might all be reasonable self-management steps that prevent a manic episode from derailing careful work one has undertaken. Ulysses contracts are a commonplace self-binding mechanism used by bipolar agents to ensure their determinations are enforced. Weiner highlighted "tentativeness" as a self-management technique. Relevant here is the *attitude* that this involves:

As invoked by the group on several occasions, tentativeness referred ... to a cultivated stance of uncertainty and suspicion toward one's own thoughts and emotions at any given moment. (Weiner 2011, 472)⁶⁹

Such an approach is likely to protect long-term projects from interruption, or from being thrown out when they begin to seem pointless. A Ulysses contract is an enforcement device that prevents a foreseen change in one's values from disrupting the pursuit of a commitment. Tentative suspicion prevents one from reasoning towards throwing out a project, and Ulysses contracts prevent one from doing so even if one reaches that conclusion. Appropriate support from friends might involve a mix of negotiation, promise-making, encouragement, and perhaps cajoling (if enough trust and insight is

⁶⁹ We have previously discussed reasons to avoid the attitude of tentativeness (§3.2.3). The danger there was that taking a pathologizing stance towards one's own view of reasons threatens to impose so much tentativeness that agency is paralysed. It does not follow from this that there is no safe dose of tentativeness, and so we should remain open to the possibility that this can be a suitable self-management strategy with respect to self-determination, under the right conditions.

present). Conceivably, this kind of self-management might attempt to harness manic flurries of activity, in order to get some commitment across the line, (however ‘feeding the beast’ is a risky strategy).

Self-management for wide-scope self-determination ensures that an agent’s affective episodes do not upset or interfere with the pursuit and eventual fulfilment of commitments. Self-determination here involves commitments being seen through. To rule oneself successfully, rulings must eventually meet the world in one’s actions. Commitment to some project closes practical deliberation, and the determination is made good when the commitment is later enforced. Given the expectation that one will not want to enact one’s commitment come the time, wide-scope self-determination tools such as Ulysses Contracts support self-determination by enforcing commitments at the relevant time.

§5.3.2 Self-management for Narrow-Scope Self-Determination

This approach secures self-determination by ensuring that the projects we are embarked on are always ones that accord with the reasons as we see them. Our practical lives are unfolding in accordance with the determinations we make, so we are self-determining. Our (perhaps episodic) values establish the rulings we make about what standards of success or appropriateness apply. The complication that BoRD presents is that these rules change with our contemporaneous experience of value. The existential reality of BoRD challenges one to self-determine despite a shifting, predictably-unpredictable, view of the normative landscape.

This kind of self-management is more deferential to an agent’s contemporaneous evaluations. The attitude of tentativeness which supported wide-scope self-determination has a counterpart in “the literal practice of avoiding commitment to future plans that would require a reliable, continuous self” (Weiner 2011, 472). Adopting this sort of regime suggests making commitments that are fairly flexible, that are less dependent on certain tasks being done at precise times, and which are unlikely to deteriorate if they lie fallow for a period. Portfolio working and flexible work hours are useful features of this kind of self-management. Support is likely to involve cultivating relationships that can survive periods of radio silence, and which have space for sometimes volatile emotions to be expressed without

being followed by recrimination.⁷⁰ This approach to self-management avoids the restrictive features of wide-scope self-determination support, as those could cost us opportunities to make new determinations that accord with our new, and perhaps episodic, values.

At the narrow scope, we do things in a self-determining way, rather than doing precisely those things that were previously determined by the self. We should readily acknowledge this sort of self-determination as a candidate to express autonomy. As in §5.2, episodic values might speak for us, for at least as long as they last, in the way Jaworska outlined. Two more conditions remain for minimal autonomy on her account. The first is met as these cares or episodic values present considerations to us as reasons to act: we self-govern when we act on them, (Jaworska 2008, 93). The final, mental freedom, condition requires that we take some consideration “as a reason against the background of the possibility of reflection”, (96). The worry is that people with BoRD are “in the grip” of their episodic values, and so unable to properly reflect. Whilst this might sometimes be the case, it is not an essential feature of depressive or manic moods as we have understood them. On Ratcliffe’s view (2010; 2013, 2020), depressions are characterized by particular existential feelings which are pre-intentional, but which scaffold and confine the normative descriptions under which the world is intelligible to an agent, and therefore, the set of normative outlooks that are available to them. As we are always in one mood or another, or having our phenomenal experience scaffolded by one or another existential feeling, the mere fact that some normative outlooks are not available to us for reflection does not render us mentally unfree. The fact that someone is in a hopeless or nihilistic depression will prevent reflection from *certain* outlooks, but on its own this does not remove the possibility of reflection or defeat mental freedom. When a person’s cares or world-disclosing affect seem foreign to another party, it can be hard to recognise reflection (or the openness to it) when it is present, but neurotypical agents are similarly restricted in the normative outlooks they can reflect on, as their more comfortable existential feelings often do not facilitate depressive nihilism. Such an episodic normative outlook can still speak for an agent, and can underpin autonomous action. Self-determination predicated at the narrow scope is

⁷⁰ Brandenburg (2017) has sketched what a nurturing stance, adopted by clinicians and caregivers towards psychiatric patients, looks like, which is a useful starting place when thinking about how interpersonal relationships might deal with blame, reproach, and supporting agents with BoRD in various aspects of life. We will discuss responsibility practices in fuller detail in Chapter 6.

deferential to one's contemporaneous values, and amounts to at least minimal autonomy as defended by Jaworska.⁷¹

The point is not that there are certain cases when we ought to prefer narrow-scope self-determination to wide-scope (however I will outline one below). Failure to recognise this sort of self-determination is undesirable as it confines the possible shapes of an agent's autonomy. Non-recognition both means that we are unlikely to be supported in our narrow-scope self-determination, and that neither we nor our peers are likely to notice that this is a space in which our autonomy could extend. Exercises of our agential capacities in that space, therefore, will not be recognised as candidate material for autonomy, and so the social recognition that autonomy relies on is absent. This means that fewer ways to be autonomous are available to us. Insofar as our concern for agents' autonomy is rooted in a desire that they be authors of the important features of their lives, the shape of their autonomy seems to be a feature they ought to have a good deal of freedom with. This freedom, if it is to be valuable, requires both that agents understand their options, and enjoy the support required to enjoy the kind of autonomy they decide to establish. Particularly when one's agential project and autonomy must confront a challenging existential reality, options to establish an autonomy of different shapes that might better accommodate that reality, are desirable. Establishing the shape of one's autonomy is, in essence, fixing the description under which one's agency is to be taken, and thereby establishing what are the conditions for excelling with that kind of agency.

Yet, consider what might be a case of narrow-scope self-determination *par excellence*. Reid (2012) writes that she "choos[es] fiercely to live—but only for now." This choice, she tells us, "sustains me on the days when I make the mistake of looking too far into the future." Living with extensive paralysis, Reid does not describe a commitment never to pursue her death, and neither should we understand her

⁷¹ Recall that in Chapters 1 and 2 we established that our emotional systems continue to be reasons-responsive insofar as they track reasons and maintain regulative guidance, even during futility thinking or the urgent restructuring of agency involved in mania. Regulative guidance and mental freedom perform the same function in these accounts, ensuring that our view of reasons and action remain imputable to us as agents, rather than to some system or phenomenon that we are not practically identified with. The background possibility of reflection is secured, even if the reflection is not going to happen, by the same attention to the appropriateness of our normative outlook that depressed and manic agents seem to exhibit. As long as we maintain regulative guidance, therefore, we retain this background possibility of reflection and so are mentally free.

as establishing a trigger for pursuing her death.⁷² What would wide-scope self-determination offer Reid? A commitment never to pursue her death leaves her a hostage to fortune, bound by a decision she might come to deeply regret should her circumstances change. A commitment to pursue her death once some trigger has occurred sees her living under the Sword of Damocles. Her circumstances could, any day, come to meet that description yet feel much more livable than she anticipated. She should then abandon that commitment, making it unclear why it was a commitment at all rather than simply a prediction about her future preferences. Reid tells us that she considers suicide every morning and resolves against it. Her commitment to live is continually tested against her contemporaneous values, and continually affirmed. Rather than having determined to live, she lives in this self-determining way.

§5.3.3 Overview

This section has outlined, in broad terms, what two self-management regimes look like and how they predicate the self-determination they support. The next two sections will explore in more detail what self-determination looks like at these scopes, before §5.6 draws out the implications this contrast has for agents living with BoRD, as well as for our thinking about autonomy and its dimensions.

§5.4 Wide-Scope Self-Determination

Self-determination predicated widely is the familiar kind, and self-management that supports wide-scope self-determination prioritises the eventual enacting of commitments, and is obstructed by affective episodes which introduce contrary values. When a commitment is adopted, practical deliberation on the matter is closed, and so a commitment involves a resolution not to deliberate further on the matter (unless new information is found). Commitments allow us to store practical

⁷² She writes, “Because I choose, fiercely, to live for the people who love me; and will continue to do so until such point as they understand I cannot carry on. I hope that moment, if or when it comes, is many years away’ which could be understood as establishing a trigger, but I think this is to misread the passage. This is, rather, an acknowledgement that a future in which she pursues her death is imaginable, if imaginable specifically under this kind of description.

conclusions for the future, when deliberating on what to do might not be convenient, (Bratman 1999). An agent can determine what they value, what they will do, and ultimately who they will be because they can rely on a tool that will cause their commitments to be effective when they are supposed to be, such as a Ulysses Contract. Radoilska (2012b) presents this as archetypal self-determination. It is valuable specifically because it protects us from obstacles to our agency which we can foresee and manage.

The advantages of such an approach are apparent. Agents can rely on their future selves to enact commitments made now and can be relied upon by others to do as they said they would. They avoid the collapse of a project they have put a great deal of work into, because in the last months a depression prevented them from completing it. Tentativeness towards their episodic values protects their long-term plans from being overturned for the sake of starting new projects that one will shortly abandon. Two limitations matter for this regime. Securing self-determination in this way costs us a degree of wholeheartedness, and predicating it this way imposes a degree of fragmentation on our self-narrative. These might be costs worth enduring for the sake of making our determinations ultimately effective at the wide scope, but we should understand how such an approach alters the shape and texture of our autonomy.

§5.4.1 - Wholeheartedness

Few people will go without ever being conflicted about what they are to do or being committed to do something they don't want to do, but BoRD threatens a more serious loss of wholeheartedness. The self-management outlined here pushes one to do something that, due to one's experience of value, is specifically not what one sees oneself as having winning reasons to do. Such changes are priced into living with BoRD, and so self-management that preserves these self-determined commitments requires that one act unwholeheartedly some of the time. Acting contrary to the best (or sufficiently good) reasons as one sees them is not the same as local losses of motivation or even doubting the value of one's projects. Affective episodes involve an experience of value that not only raise doubts about the value of prior commitments, but frequently suggest that those commitments are unjustified. Self-management that presses one to stick with commitments will help protect one's overall agential project

from upset, but it ties one into localised irrationality during affective episodes, when one must act unwholeheartedly.

This unwholeheartedness costs us more than just comfort. It costs us some of the *self-answerability* for our commitments that marks a non-passive relationship to our values, (Westlund 2009). Westlund argues that if we lack this disposition to give account for our normative outlooks, we lack an attitude that is crucial to our mental freedom. Whilst we might be able to *recite* the reasons we had when we made a commitment, if they are contrary to the reasons as we see them now then we are not exhibiting self-answerability when we act on them. Acting on commitments despite “considerations that purport to challenge our current sense of the justificatory landscape”, involves abdicating self-answerability for those commitments, raising the threat of “passivity in the face of one’s commitments” (34). On the contrary, manic agents who wax lyrical about the merits of their exciting new projects, as well as depressed people who offer articulate nihilistic outlooks, are exhibiting self-answerability. Their interlocutors might not be persuaded by the accounts they give, but their disposition to offer such accounts demonstrates the relevant attitude. Agents who do as they resolved to long ago might exhibit strong wills, but if this comes at the expense of acting on the reasons as they see them now; the agent is not *ruling* themselves now. The ongoing practice of determining what one will do is given up. This is where scope begins to matter, as predicating self-determination at the wide-scope involves *unpredicating* it at the narrow-scope.

§5.4.2 - Fragmentation

Composing a self-narrative involves making the things we do and experience collectively intelligible, and a self-narrative is robust to the extent that it makes more of our agential life intelligible. These narratives pick out the agent who does these things, because they see the world this way. As we saw in the previous chapter, a self-narrative, or some part of a narrative, might become empathically inaccessible to us if our normative outlook shifts due to an oncoming depression or manic episode (see §4.3). A life united by the same, or substantially compatible, overarching projects is a life that will be easily narrated into a cohesive whole in certain respects, but where securing these projects imposes the kind of unwholeheartedness noted above, it threatens to fragment the narrative in another respect.

The usual worry for autonomy is that the things we do during affective episodes will not cohere with our usual self-narratives (Potter 2013) or with whichever are our identity-constituting considerations (Willigenburg & Delaere 2005). It is reasonable, in the face of a single depression which interrupts what is otherwise a well-integrated and intelligible life, to jettison it—to decide that it *should not* be accommodated by one’s self-narrative. If the story hangs together better, allows one to confront the world more effectively, and allows others to participate in reliable and valuable relationships with one, if that period of one’s life is consigned to unintelligibility, then casting it as an interruption to the narrative rather than part of it might be appropriate.⁷³

However, the shifting evaluative sands of BoRD are often core to the experience, and many agents (such as James and perhaps Mr. M) are not able to compose a self-narrative based on any normative outlook which is independent of BoRD, and if they did, any such narrative would be interrupted by the outlooks that BoRD imposes. Many people with BoRD encounter episodic values, and revert to any longitudinal values, many times over in their life. Strategies which secure wide-scope self-determination ensure that we do things which can be accommodated by some self-narrative which is abstracted from our symptomatic experience of value, but that identity surrenders significant portions of our experience to unintelligibility. The agent does not recognise the narrative as making sense or as successfully picking them out. Agents who do possess longitudinal values might be able to compose abstracted self-narratives, and so would be able to cast their manias and depressions as interruptions to that identity. Yet, much more of their life is therefore composed of ‘interruption’. Sense cannot be made at once of the things they do during these times and the way the world seems to them. Conspicuously, this is the move that Jamison avoids when she concluded that she is probably both her mania, depression “and much that is neither.” (1995, 93). Jamison could choose a narrative that abjured these episodes, but instead she acknowledges that they make her up as much as do her other experiences and projects.

⁷³ Refusing to make *any* sense of such episodes is an uncommon reaction, but a more acceptable way of jettisoning these parts of our lives is by placing them on the predicament side of the protagonist-predicament distinction, (Schroeder 2022, 219). Reifying depression into a medical phenomenon that happens to an agent makes it a given part of the world (predicament), rather than an experience of the world that has standing *inter alia* with one’s more familiar and desirable experiences of the world, and so demands intelligibility.

Enforcement devices like Ulysses Contracts protect commitments but they do not make one see, in the moment of enforcement, the reasons behind the commitment. We can look back on the time they were enforced and see that what we did cohered with a narrative we endorsed at the time we made the commitments, and perhaps also at the time of reflection, but we cannot, make sense of how it felt in the moment of enforcement. Enforcing commitments secures effective wide-scope self-determination, but severs the experience of acting from the contemporaneous experience of value. These are both elements that would be united in more robust narratives. This kind of self-determination commits one to a particular degree of fragmentation in self-narrative, as the normative outlooks which run against protected commitments are shut out of one's practical identity. Large swathes of experiential life are not to be integrated. Securing self-determination, predicated widely, tugs at the integrity of one's self-narrative.

Securing self-determination, predicated widely, is familiar and deeply valuable to many people living with BoRD. This examination, however, reveals where key trade-offs are made when we secure wide-scope self-determination. Self-management which secures this wide-scope self-determination is not *gratis*, and room must be made for it at the expense of other dimensions or features of autonomy. If agents with BoRD are to craft self-management regimes to manage their symptoms, it is important that they know what effect securing self-determination, predicated at this or that scope, will have on the shape of their overall autonomy.

§5.5 Narrow Scope Self-Determination

Self-management which secures self-determination at the narrow-scope is likely to prioritise resilience, flexibility, and insurance policies. This might ensure that the commitments one makes are to projects which can be effectively put on hold, perhaps for weeks at a time, and which can be moved between with relative ease and minimal cost. Artistic projects can often be left fallow and picked up again without having deteriorated. Part-time or folio working might be preferred. Self-determining in this way

ensures that agents are always engaged in the project of determining what they should do, and are fully present in their actions.

Jaworska has previously (1999) argued that a contemporaneous capacity to value is all that need persist for a person to be owed deference in virtue of their autonomy. It is specifically in the self-determination dimension that this is important, as it is the capacity for the agent to value that allows them to determine what is valuable *for them*, and to determine what makes their life go well. As these determinations are upstream of the agent's commitments to act in certain ways, narrow-scope self-management protects self-determination by ensuring that one's commitments stand in proper relation to one's values, even when that requires that one revise some commitments. Passivity, or 'doing the done thing' is the quintessential failing of self-determination. An agent who acts in this way doesn't determine for themselves what is valuable, and does not place their own values in the proper relation to their behaviour. These activities are core to the business of being an agent, and at whatever degree of specificity we deploy the description 'agent', we place an expectation on the agent that they engage in these activities. To live an acquiescent life, passive in the determining of what one does, is to fail at being whatever kind of agent one is. To live an acquiescent life is to suppress crucial agential capacities, which cannot be the foundation for our autonomy unless they are exercised. It also too readily invites explanatory answers to the question *why* one did what one did. The explanation is that one was asked or instructed – the agency of those from who one takes instructions is what we must take an agentic stance towards. We can even be passive in the face of our previous commitments. Enforcing commitments that we no longer see to be upheld by good reasons secures wide-scope self-determination but gives up our place as the constant arbiter of what we shall do. If an agent were to have serious worries that their present commitments were not underpinned by good reasons, yet fulfilled their commitments without exploring those worries, they would invite criticism for this passivity (again, Westlund 2009). Self-determination predicated narrowly involves an ongoing affirmation of one's values and of one's projects.⁷⁴

⁷⁴ Acting on second thoughts is a common and important exercise of autonomy (Radden 1994).

Several advantages of narrow-scope self-determination are worth noting. Agents living with BoRD are better able to develop an integrated practical identity that can ride stormy normative waters without bits falling off. Agents are protected from feelings that their plans are wrongheaded or wasteful of their lives. This inner conflict is not so egregious that we should *always* prefer not to experience it, but it is often worth avoiding. Agents are answerable for what they do in a way that is sometimes lost when self-determination is secured at the wide scope. Self-determination predicated narrowly sees one *ruling* oneself, rather than issuing and enforcing individual rulings. Self-determination, here, is an ongoing project of affirming one's values and commitments. Rather than telling us what list of deeds an agent has determined for themselves that they will do, this tells us when an agent does something in a manner that secures non-passivity. Wide-scope self-management protects us in doing the things that we determined we should do, whereas narrow-scope self-management protects us in doing things in this self-determining way.⁷⁵ Wide-scope self-determination protects rulings, and narrow-scope self-determination protects the practice of ruling. The relative degrees to which we secure wide and narrow scope self-determination will greatly alter what our agential project looks like, and make a difference for the shape and texture of our autonomy.

There are, of course, limits to narrow-scope self-determination. Many commitments are less likely to be effective, as deference to our contemporary evaluations makes revision of commitments easier. Being diverted at the crucial moment might cost us a project we have spent years on, meaning not only that certain commitments are made ineffective, but that a large chunk of our agential project also becomes ineffective.⁷⁶ If this happens, we fail to self-determine at the wide-scope. Self-management which prioritises ongoing self-determination, in this narrow-scope sense, endangers the determinations we made before these episodic values emerged.

⁷⁵ I have in mind something like the distinction Constantine Sandis explores between an action and the *way* the action is done, (2015, §2). Self-determination might be a thing we do when we make particular rulings about what are our values and projects, or it might be the way in which we conduct agential life. (Compare, Steward 2012).

⁷⁶ I mean two different things by 'effective' here. Our projects are effective or otherwise depending on whether we successfully bring them to fruition. Our agential project is effective or otherwise depending on whether it eventually determines what we (try to) do. If our projects are unsuccessful, for reasons other than revision or abandonment, our commitments and *agential* project has still been effective.

§5.5.1 - An Effective Agential Project

We can exaggerate this worry, however. Holton (1999) argues that we should identify weakness of will specifically with a tendency to revise our commitments in cases where we should not, and takes “contrary inclination defeating” intentions to be the commitments which, when revised, are most likely to make a charge of weak will appropriate (248). These are the intentions we form with the expectation that we will, in the future, be less certain of but which we hope to still be motivated by nonetheless.⁷⁷ The ‘should not’ standard, here concerns a particular sort of practical rationality. Tendencies to revise where we should not are those tendencies that, overall, endanger the efficacy of our agential project. Citing Bratman, he argues that a weak will is one that exhibits such a tendency to revise commitments too often, (Bratman 1999, 64, in Holton, 247). Whilst I grant that narrow-scope self-management costs us some wide-scope self-determination, I do not grant that it invites the charge of weak will.

The standard of appropriate revision is set by the requirements of an effective agential project, yet it is not clear that this standard is non-agent-relative. The existential reality of BoRD involves highly varied and changeable experiences of value, which are likely to cost the agent some efficacy in their overall agential project. However, part of what an agential project does is confront our existential reality, making it both intelligible and livable. We should expect some differences in what these projects look like when confronting starkly different existential realities. To illustrate this point, consider what an agential project in the face of BoRD would look like were the revision standards to be the same as those for neurotypical agents. We would discover two problems.

Calhoun (2009) describes *demoralization* as a way in which we fail to behave as agents without failing to count as agents. She tells us that:

“The possibility of agency ... depends not only on being moved by one’s deliberation, but also on being moved to deliberate in the first place. The point of deliberation is to affect the world through one’s actions (or inactions). Under normal conditions, we take for granted that deliberation has a point: Our actions do affect, or stand a good chance of affecting, the world in the ways we intend. Most of us, most

⁷⁷ Parfit’s Russian nobleman (1984) has a ‘contrary inclination defeating’ intention to bequeath his land to his serfs, and builds insurance into his will to protect against contrary inclinations. The intentions we protect with Ulysses contracts or other wide-scope self-management techniques are of this kind.

of the time, simply take for granted the practical efficacy of our deliberating and acting. But that background frame can be disrupted. When it is, agents experience a demoralizing loss of confidence in their ability to function as an agent.”

(204)

If the standard for reasonable revision of commitments is the same for agents with BoRD as for those without, the loss of efficacy in one’s agency makes this sort of disruption more likely. Changes in one’s experiences of value sometimes lead one to the conclusion that what one is doing is not worthwhile. To the extent that these conclusions become pervasive, the only way for the agent to avoid a weak-willed revision of their commitments is to continue doing things they see themselves as not having particularly good reasons to do. Their agential project becomes increasingly composed of things that are ‘not worth doing’ and, whilst it might still be worth having *an* agential project, it is hard to see why the agent should be invested in *this one*.

The second problem is that, a well-crafted self-management regiment notwithstanding, BoRD will still make us less effective in at least some of our commitments. If we take the revision threshold to be the same for someone living with BoRD as for someone not, the person living with BoRD will frequently not have leave to revise their commitments, and so more of their commitments will fail rather than be revised. To avoid the charge of weak-will, someone living with BoRD must endure higher rates of failure in the projects they adopt. This builds a dataset that will eventually suggest that one’s agential project is not working. The rational conclusion for the agent to draw from this evidence is that practical deliberation does not produce an eventual effect on the world, and so the standards for an effective agential project *drive* a demoralized breakdown of agency.⁷⁸ This is the second path to a collapse of agency produced by a non-agent-relative standard. Steps taken to protect the efficacy of one’s agential project should not, I contend, eviscerate self-trust and make one’s agential project seem futile. Eventually, the requirement that one revise commitments only when it would be appropriate for a neurotypical agent amounts to a requirement that one live as if one did not have BoRD, and for a given severity of BoRD such a requirement is perverse.

⁷⁸ Such a conclusion, rational or otherwise, is not one of the particular kinds of self-distrust that McLeod and Sherwin outlined (2000, see §4.2 above). It is a deeper sort of distrust in the practice of deliberating, which casts agency as not only unreliable (type 2 distrust) but as ineffective and detached from the world.

Self-determining by revising our commitments is not some sort of creative accounting: it is an important part of managing our agential projects in the face of our practical experience. Holton's argument highlights that a degree of efficacy in our commitments is the cost for securing self-determination predicated at the narrow-scope. What this does not entail is that narrow-scope self-determination is a fresh coat of paint on a weak will. This is another sort of self-determination that is available to people living with BoRD and what it can provide is valuable. It allows one to craft an agential project that, for all its abnormal shape and texture, is better suited to the existential reality one faces, and the challenges peculiar to that reality.

§5.6 - What Does This Tell Us?

So far I have argued that understanding self-determination as a self-rule concept predicated at particular scopes clarifies the choices one must make if one is to support autonomy in this dimension whilst living with BoRD. Previously we have understood the ways in which autonomy extends more or less far in various dimensions to be what gives bipolar autonomy its peculiar, lumpy, shape. What consequences does our understanding of self-determination as coming predicated at different scopes have for the shape and texture of bipolar autonomy? How should we describe the kind of agency that is available to agents living with BoRD, and what ideal of autonomy is local to their kind of agency?

§5.6.1 The Shape of Autonomy

I have argued that there are two distinct ways in which adopting suitable self-management regimes allows bipolar agents might support self-determination in the face of the existential reality of their condition. These regiments, and their differences, are simplified in order to highlight their contrast. In practice, I, and I expect most people living with BoRD, craft self-management regimes that secure self-determination at these scopes to different degrees, rather than prioritising one. The right balance will depend on a range of circumstances, including which dimensions of autonomy are most important to an

individual, as well as what shapes of autonomy are realistically available to them. Recall that different agents' autonomies will vary drastically in their shape, how far it extends in different dimensions, (Killmister 2017).

Bipolar agents can choose what kind of support to put in place for their self-determination, and to what degree they want to secure self-determination at particular scopes. Examining options and their implications begins to tell a richer story about the shape and texture of an agent's autonomy in general, and its self-determination dimension in particular. BoRD reshapes agents' autonomy, but not in a monolithic way. As we have seen, BoRD creates relations between different dimensions, some of which are inverse relations. Greater widely predicated self-determination runs inverse to whichever dimensions are injured by fragmentation of one's self-narrative.⁷⁹ Rather than imposing particular shapes on autonomy, BoRD confines the range of possible shapes an agent's autonomy might take. There are limits on how far autonomy might extend in this or that dimension, and extending further in one dimension drags the frontier back in another. Understanding these restrictions on possible shapes helps us to better specify the kinds of agency available to people living with BoRD. This is not just whatever autonomy that is left unobstructed by BoRD, but a narrower description of what autonomy looks like when a person living with BoRD cultivates some particular agential project. Understanding this possibility space reveals something about the kind(s) of autonomy that can be ideals worth striving for from the bipolar agent's starting point.

Agents who favour wide-scope self-determination are much more likely to see their commitments eventually come to fruition. Effectively determining what one does has practical and normative consequences for one's future, and allows one to craft a life where projects are more stable and reliable and where friends can better rely on one, but where unwholeheartedness often looms and one's self-narrative can become fragmented. One must contend with experiences of value that are not intelligible together with the rest of agential life. Narrow-scope self-management secures self-determination as a continuous process of affirming what one does, allowing the agent to be fully present in their actions. Whilst this costs a degree of efficacy, it need not cost so much that we need worry about weakness of

⁷⁹ Examples might be self-unification or perhaps self-definition, on Killmister's theory (2017), or self-governance on Mackenzie's theory (2014, compare Korsgaard 1996).

will. In cases where it does weaken an agent's will, that agent has reasons to adjust the balance in which their self-management is supporting wide and narrow scope self-determination. Understanding self-determination at scope allows us to better inform people living with BoRD of the options they have, and how they might secure autonomy of some particular shape that suits their life. Fixing the scope at which one intends to secure self-determination, in whichever domain of life, more narrowly specifies the kind of agency the particular agent is cultivating. This is stipulated within limits set by one's existential reality, but also within a space of relative freedom in which we get to decide which kind of agent we want to be – and therefore what the local ideal of autonomy looks like for us. Specifying this kind of agency allows agents to take control of the agential project they pursue, and set for themselves the success criteria that apply to their project. Such stipulation must be a crucial exercise of autonomy, if autonomy involves being the (co-)author of one's own life.

In previous chapters we have established that bipolar autonomy is available, even if it is lumpy and not immediately recognisable to agents who are used to very neurotypical kinds of autonomy. There are, therefore, a great many more ways to be autonomous than has previously been recognised. What we have established here is that within the realm of marginalised, neuro-atypical sorts of autonomy, there is great diversity again. This diversity is not always a given fact of agents' situations, the way one's existential reality is. Some of this diversity is the result of different people responding to those existential realities by establishing different kinds of agential project, which are each the foundations of autonomy of a very different shape. All agents must establish the shape of their autonomy within whatever confines are set by their existential reality, and bipolar agents must develop their autonomy into a shape that accommodates the particular experiences and challenges they will face. Which agential capacities they wish to develop, in which domains, and where they hope to enjoy high degrees of recognition, are decisions as to the shape their autonomy should take.

§5.6.2 - The Texture of Autonomy

In order to be informed when making these decisions, agents must be sensitive to one further distinction. Self-determination can be predicated at different scopes and that at these scopes it will require different conditions; its exercise will look markedly different; and it will produce different shapes in one's autonomy. Self-determination, at these different scopes, still does the same work of ensuring

that an agent is not living a passive or acquiescent life. Non-passivity is secured by wide-scope self-determination as when agents make rulings as to what one will do and then executes them, they are exercising their agency and avoiding the arrest of agency that James exhibited (Weiner 2011, see §3.2.3). At the narrow scope, non-passivity involves a continuing affirmation of one's commitments and (perhaps episodic) values, involving an openness to revise commitments if one sees good reason to. Supplementing Killmister's understanding, I have referred to the shape *and texture* of agent's autonomy throughout. The distinction in scope put forward in this chapter offers us the conceptual apparatus to track these textural differences. We can speak of someone being more or less autonomous, as well as their autonomy extending more or less far in various dimensions, but we need to also heed the aspects of autonomy that are decidedly more qualitative.

In this chapter I have argued that self-determination can be predicated at wider or narrower scopes. If one's autonomy cannot be extended far in this dimension when predicated widely, we might decide to craft an agential project that secures self-determination predicated more narrowly, so that we still enjoy some self-determination. By maintaining this shape of our autonomy, we are making a qualitative, rather than extensional, difference to our autonomy (at least as it extends into self-determination). This is a textural difference. An agent's autonomy might extend significantly in the self-determination dimension, but this extension will be 'porous' if it is predicated widely at the expense of narrowly predicated self-determination. It extends in this dimension, but when we zoom in to look very closely we can notice all the gaps where they are, in some moments, passive with respect to their view of reasons, trusting to their self-management practices rather than to their sense of what good reasons are at that time. Alternatively, significant extension in this dimension might occur when one self-determines at a narrow-scope, but this extension might be 'thin' or 'fuzzy' or 'viscous' – depending on how we want to describe the property of not quite constituting a mereologically contiguous body. Our self-determination comes in substance rather than dry objects. We might yet find better vocabulary for these variations, but as long as it has clear referents, this account will allow us to tell rich stories about the texture of agents' autonomy. These textural differences also play a role in specifying what kind of agential project one is embarked on. Some agents might prefer to rely on sturdy and robust self-determination, and so secure it at the wide scope, allowing discreet determinations to bridge all the gaps and hold the structure together. Another agent might prefer the freedom to *swim* through their agential life, and so secure self-determination at the narrow scope to preserve maximum freedom to

shift and swerve and move around in this dimension of autonomy. Agents who are distressed by uncertainty, who rely on the predictable, and for whom pattern and habit are important, might more easily belong to the former camp, whereas the ardently spontaneous, easily bored, and hyperactive might find the latter more suitable. These are choices that remain open to agents who are crafting an agential project, and so are another avenue for agents to stipulate for themselves the kind of agency that they want to have, and consequently what the ideal of autonomy looks like in their locale.

§5.6.3 - Recognition, and the Social Context of Self-Management

To the extent that friends, family, and clinicians wish to help agents strive for an ideal of autonomy, understanding the various paths to autonomy of different shapes will reveal more ways to help agents achieve that ideal. Understanding where our peer's autonomy extends and does not extend is crucial if we are to confer appropriate recognition on it, but sensitivity to the texture also matters. Recall, these differences serve to more narrowly specify the description 'agent' according to which we set expectations on someone to excel as an agent. The kind of recognition we owe is recognition of the agent under a fitting description, so to the extent that these textural differences specify the fitting description, they partly determine the kind of recognition that is owed.

If social recognition of one's agential capacities being at work is necessary for agency to be autonomous, as we assume, then the exercises of agency involved in stipulating the kind of agency that one is embarked on requires recognition. To properly respect an agent as autonomous, we must recognise them as an agent under the description they have stipulated.⁸⁰ For the exercise of these stipulating capacities to become autonomous, one thing that is required is social recognition of those capacities when they are at work. This is to say, the shape and texture of an agent's autonomy is counted as their achievement, rather than their reality or predicament, only when we recognise the agency they exercised in producing those shapes and textures. This recognition has downstream consequences for the rest of our agency. Autonomy is socially situated, in part, because effective agency is always

⁸⁰ It might be objected that we should not recognise people as highly successful and authoritative agents simply because they say so, if they are in fact ineffective and arbitrary. To stipulate, as I use the term here, is a success verb. We stipulate the suitable description of our agency only when we successfully craft an agential project that fits that description. Unlike self-narrative, in which we can be radically mistaken, we stipulate the kind of agency we are embarked on only by embarking on it.

scaffolded by various features of the material and social world. Virtue is scaffolded by the positive reinforcement that comes from praise and flourishing relationships. Attention is scaffolded by online sorting algorithms. For many bipolar agents, effective agential projects are scaffolded by the self-management practices discussed here, which involve the support and participation of friends, family, clinicians, institutions, and (to perhaps a very minimal degree) strangers. In order to play their roles in the self-management practices that scaffold self-determination, predicated widely or narrowly, these people must recognise where a bipolar person's agency extends and where it does not. They need to know what counts as their agency and what is a failure of that agency if they are to confer the right sort of recognition.

This is particularly important for agents who rely on self-management regimes. It is perfectly likely that some agents will decide to shift to a quite different kind of self-management regime from the one they have cultivated, perhaps striking a drastically different balance of wide and narrow-scope self-determination. Extracting oneself from a regime that isn't working will often look like failing to practice self-management as it was agreed with friends, family, and clinicians. For this reason, it is important that peers are aware of the range of shapes autonomy might take, and recognise the agency exercised when reshaping one's agential project mid-journey. Awareness of the diversity in autonomy, and recognition of these shaping choices, allows those people who play a role in scaffolding our agency to distinguish changes from failures. There is a limit to how often we can drastically alter what we take self-management to look like without this relationship failing to be effective, but within those bounds, space for change in response to evidence is important.

§5.7 Conclusion

My two aims in this chapter have been to consider how bipolar agents might secure self-determination in the face of the particular challenges BoRD presents, and to illustrate the range of choices agents (bipolar or otherwise) face when crafting autonomy with some particular shape and texture. The scope at which a bipolar agent wishes to secure their self-determination will alter the kind of self-management

they should practice, and maintaining such a regime will produce a life marked by quite distinct sorts of self-determination. Predicating at scope lets us capture the practical dilemmas faced by people living with BoRD, and is useful for deriving a rich and informative story about individuals' autonomy. Some of the more arcane differences between self-determining exercises or lives can be explained by specifying the scopes at which we predicate self-determination in each case. Understanding that the choice to predicate self-determination at this or that scope tugs at the fabric of one's autonomy in other dimensions opens up a range of discussions about how autonomy might be crafted to better confront different existential realities.

I suspect that attention to the scope at which exercises of autonomy are carried out will prove instructive in other dimensions. Our attention to self-narratives in the previous chapter was concerned with how they secure self-trust and scaffold effective and autonomous agency, however as we noted in §5.4.2, self-narratives play important roles in other dimensions of autonomy on various theories. The choice whether to accept a degree of fragmentation in one's self-narrative for the sake of wide-scope self-determination was a trade-off available to bipolar agents. Might our self-narratives also admit of scope? If they do, understanding the difference narrating oneself at a wider or narrower scope would make for the shape or texture of our autonomy is also valuable to agents if they are to choose to craft a suitable autonomy for themselves. Given that we cannot easily abandon our pasts if we are to be autonomous or maintain valuable relationships, any difference in the scope at which we narrate a self will likely be situational for an agent, more often than they will track differences between agents' existential realities.⁸¹ We cannot pursue this line of thought in any further detail here, but it remains a valuable avenue for further research.

We have established that bipolar autonomy comes in lumpy varieties that are often harder to recognise for those used to a neurotypical standard. Exercises of this autonomy can go unnoticed, or can be

⁸¹ However, we can surely imagine some agents who have powerful reasons to leave their pasts behind, and who should be empowered to do so. England's aristocracy relied on Prince Hal 'turning away his former self' for the sake of avoiding further civil war and pursuing the projects they valued, (conquering France). The character *Piscine* from Martel's novel *Life of Pi* seems to be offering a fairly revisionary narrative of how he survived a shipwreck, but the reader is left sympathetic to his plight, and understand why the fictional investigators make his version of events the official account.

mistaken for failed exercises, due to the bipolar agent's interlocutor's unfamiliarity with this kind of agency. Self-determination predicated narrowly often looks like failure of self-determination predicated widely. It is unsurprising, then, that bipolar autonomy is widely endangered by non-recognition, which undermines the basic social conditions of autonomy. Considering this risk, we must ask how this recognition can be secured. One solution is to consider the act-descriptions in play when bipolar agents exercise their lumpy autonomy. Friends and family who play supportive roles in a bipolar agent's self-management strategy are, *ex officio*, primed to describe some actions as failures of self-determination. However, an act-description that will sometimes be appropriate in these contexts is that the bipolar agent is shifting to another self-management regime, striking a new balance between wide and narrow scope self-determination. The bipolar agent can reassure their intimates that they are, in fact, exercising their autonomy by offering this act description. Contesting and negotiating act-descriptions is a valuable exercise in drawing other's attention to our view of reasons, and our agential capacities when they are at work. In the final chapter I will argue that responsibility practices, at their best, prompt this kind of exchange and so offer a valuable and ubiquitous method of securing greater recognition of agential capacities, which better satisfies the basic social conditions of autonomy.

Chapter 6 - Acts, Descriptions, and Responsibility as Answerability in Mania

6.1 Introduction

Chapters 1 and 2 examined agential capacities to track and respond to reasons as they are found in common sorts of mania and depression. They are often limited, but not impaired beyond use in depression, and are pervasively restructured in mania in ways that impose limitations, but which also offer situational advantages. The driving thought of this inquiry is that mad people are owed recognition on their own terms, recognising of the specific kind(s) of agency available to them. There is agency here to recognise, but what remains unclear is what form this recognition should take. Recognising mad agency on its own terms involves recognising the (perhaps putative) reasons that an agent is responding to, and recognising what they do as a response. The interpretations adopted by manic people's non-manic interlocutors do not enjoy priority or default status in matters of recognition, and the manic person's preferred description of their deed offers insight into what their intention was, in the usual Anscombian way (2000, §23). However, a person's intentions are not the only thing that matter if they are to successfully live with others and enjoy interpersonal relationships that mutually facilitate flourishing. Recognition must be extended in ways that allow bipolar autonomy to be established *inter alia* with the other goods of a flourishing life.

In this chapter I investigate how responsibility practices might be virtuously carried out when the person being held responsible is manic. I focus on responsibility for two reasons. The first is because it is one of the most common instances when a person's act-description, and disagreements over act-descriptions become relevant. Insofar as recognition is necessary for autonomy, settling on appropriate act-descriptions becomes important if manic agency is to be recognised during responsibility ascriptions. I argue that virtuous holding-responsible practices confer recognition on agents' agential capacities in ways that satisfy the recognition condition for autonomy, meaning that holding an agent responsible not only respects them, but renders them (more) autonomous by satisfying the basic social conditions for autonomy. The practice of remonstrating with another person serves to draw their attention to how their behaviour has affected us. Often this involves demanding that the person we are holding responsible answer for what they have done under *our* act-description rather than theirs, but ours is not

always the best or most fitting description. The second reason is the highly intuitive link between the things we do autonomously, and the things we might be responsible for, and more generally, the global property of autonomy and the general property of being a responsible agent. On the view I outline here, these phenomena overlap insofar as they refer to, and must confer recognition on, someone's agential capacities. The practice of holding agents responsible and conferring autonomy on them stand in a cyclical, re-enforcing, relation.

§6.2 will outline how I understand responsibility, with particular attention to responsibility in marginal agents, drawing on Shoemaker's important work in this field (2015). §6.3 will present my argument for an answerability first conception of responsibility, arguing that evaluating the quality of an agents' judgement is continuous with evaluating their regard, informed by Scanlon's relationship-maintenance view (2008). This is due to the way contestable act-descriptions encode the reasons for, and against, an action. §6.4 will argue for the importance of jointly establishing act-descriptions when holding agents responsible, and outline how such negotiations must unfold when manic people are being held responsible by non-manic peers in ways that properly recognise agency. §6.5 will outline what this account of responsibility tells us about how manic people's agency must be recognised if it is to be rendered autonomous.

Throughout this enquiry, I have placed particular importance on taking the phenomenology of mania, and other sorts of madness, seriously. This is because these phenomena involve particular existential realities that an agent has no option but to confront, and because the distinctive features of these realities include the phenomenal experience of reasons. What attention to these experiences has so far revealed is that we require new conceptual tools to deal with them in philosophy of action and ethics. The same will hold true for responsibility. Consequentially, a methodological requirement of this investigation is an openness to fairly revisionary views of responsibility. These existential realities cannot help but have an effect on the kinds of relationships people who experience mania enjoy with their intimates, and so what it takes to maintain or repair such a relationship is likely to be different from that in neurotypical relationships. I work under the assumption that the maintenance and repair of interpersonal relationships is at least included among the normative justifications for holding agent's

responsible (Scanlon 2008), perhaps in addition to the norm of properly appreciating what another agent has done, (Wallace 1994).

§6.2 Marginal Responsibility

Under the reactive attitudes paradigm, to hold someone responsible is classically understood to involve holding a particular sort of stance towards an agent that disposes one to emotionally engaged *reactive attitudes* towards them or their deeds. Resentment is a species of anger that reacts particularly to how another person has behaved towards us. Per Strawson's formulation, we hold people responsible insofar as we have the right kind of attitudes towards them, which react to their agency in the right way. This contrasts with the objective stance we might take towards a person.

The objective attitude may be emotionally toned in many ways, but not in all ways: it may include repulsion or fear, ... But it cannot include the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships; it cannot include resentment, gratitude, forgiveness, anger, ... If your attitude towards someone is wholly objective, then though you may light him, you cannot quarrel with him, and though you may talk to him, even negotiate with him, you cannot reason with him.

(Strawson 2008, 9-10)

This is similar to, but not the same as, the distinction in stances that Mitova outlines when discussing the kind of injustice involved in seeking to explain rather than recognising another person's actions (2020). We empower people when we take an agentic stance towards them, seeking answers to the *why* question in terms of how the agent saw and handled reasons. To do so opens us up to a range of reactive attitudes about the agent's recognition and handling of reasons.⁸² Taking an explanatory stance can be consistent with holding the reactive attitudes that mark out our responsibility practices – we

⁸² These attitudes have conventionally been understood to necessarily be emotionally involved in the way Strawson describes, but Pickard (2011, 2017) offers a model of reactive attitudes which are fitting responses that recognise another's agency, but allow for a degree of emotional detachment. Such 'responsibility without blame' is, she argues, specifically useful in holding responsible people with borderline personality traits. More recently, Brandenburg & Strijbos (2020) have explored a responsibility stance used in psychiatric contexts that is a degree more emotionally separated than conventional interpersonal responsibility, but not wholly detached.

might explain a person's action with reference to traits or characteristics to which we think certain reactive attitudes are fitting. A person who is greedy and thoughtless behaves in certain ways *because* they have these characteristics, whilst another behaves in a kind and giving way *because* they have other characteristics. We can be emotionally engaged in resenting or being proud of our peers for having these characters. However explanatory stances frequently overlap easily with objective stances. When we adopt the objective stance, seeing others "as an object of social policy; as a subject for what, in a wide range of senses, might be called treatment" (Strawson 2008, 9) we will tend to explain that person's behaviour with reference to their mental disorder, their culture, their situation, or some other feature that does not make reference to their normative outlook or handling of reasons. We can hold people responsible whilst taking an explanatory stance towards them, but for familiar reasons to do with properly recognising agency, this stance will frequently be inappropriate. Thus, we must understand how to hold people responsible in a way that recognises their agency, even when a potentially attractive explanation is available.

§6.2.1 Shoemaker's Tripartite Theory

Shoemaker (2015) argues that three faces of responsibility come apart in cases of marginal agency: attributability, which evaluates an agent's character, answerability, which evaluates an agent's judgement, and accountability, which evaluates an agent's regard for others. I will outline each in turn.

Attributability evaluates an agent's moral character, composed of clusters of their *cares* and *commitments*. He draws on both the broadly Frankfurtian tradition for an account of cares, and Watson's broadly Platonic account of authenticity in terms of our evaluative judgements for our commitments, (Shoemaker 2015, 45). We hold people responsible in the attributability sense when we make moral evaluations of their character, which is established by the tight weave of these two strands of the deep self. The things we care about in Frankfurt's sense are often accompanied by judgements as to what is valuable, which network together into clusters that establish who we are, at least for the purposes of being held responsible.⁸³ When we admire someone, Shoemaker argues, we are evaluating

⁸³ This *ecumenical deep self theory* (Shoemaker 2015, 47) is in some ways reminiscent of the account of how our normative outlooks emerge from §1 above. On my view, our emotional reactions and judgements contribute us a picture of the normative landscape, however this picture is not something that constitutes a deep self. Normative

them as an agent, holding that they are fit for this kind of reactive attitude in virtue of the moral quality of their care-commitment clusters.

Answerability is the face that specifically evaluates an agent's judgement, construed widely to include the way we handle reasons in cool deliberation, and when we act from heated emotions.⁸⁴ To be answerable for an action is to be "open, in principle, to demands for justification regarding that [action]" (71). It can sometimes be ambiguous whether an agent is open to such demands, especially if their stated reasons or act-descriptions are not accepted by, or even intelligible to, most of their interlocutors. This sort of openness obtains when an agent has quality of judgement, which they have just if their attitudes are "generally governable by the agent's judgements about why they are more worthy than (some relevant) others" (82). In short, we can demand that an agent answer for what they have done as long as they can tell us why their reasons for acting were better than their reasons for not doing what they did. These are the reasons to do something, or nothing, instead of the action in question, and so Shoemaker refers to them as the 'instead-of reasons' (or simply the '*instead-ofs*'). Shoemaker tells us that which *instead-ofs* are relevant in any situation will be determined by a number of factors, including context and the relationship that stands between the agent and their interlocutor, (76 f16). To hold a person answerable is to evaluate how they handled their reasons, and it attaches to the practice of demanding that agents answer for what they have done.

Finally, accountability evaluates the *regard* an agent holds another in. Broadly, regard refers to our recognition of the independent moral importance of others. Ill regard can be manifested by indifference to another person or to their interests (for instance, if we are emotionally unaffected by their being hurt), or by a failure to recognise certain facts about that person as normatively salient (that is, a failure to recognise those facts as reasons, 102). More than this, failure of regard can consist in failing to take seriously another agent's *perspective* on reasons, including "[their] attitude towards treatment by others" (97). Holding someone in good regard involves taking them seriously, where "[f]or me to take you seriously, your view of the goodness or badness of [some] treatment... must count for me too *qua*

outlooks are, at any moment, the view we have as agents, but they do little work to individuate us and are not enough to fix the description under which we should appear on our side of interpersonal relationships.

⁸⁴ Shoemaker is also working in the Sentimentalist tradition (Shoemaker 2015, 19, compare §1.2 above).

practical agent, and in sync with how it counts for you” (ibid, original italics). There are two ways in which someone who talks loudly on an airplane while others are trying to sleep seems to hold his fellow passengers in improper regard. The first is his failure to recognise facts about them as normatively salient (that they are trying to sleep is a reason for him not to speak loudly).⁸⁵ The second is when he recognises the fact that others are trying to sleep as a reason not to speak loudly, but judges that it is outweighed by other practical reasons he has, (98). We are appropriately angry towards people when we recognise that they do not take us seriously or take facts about us to have a bearing on how they should conduct themselves.

In §6.3 I will argue that responsibility trades in answerability and that answerability, when properly understood, incorporates an evaluation of regard. Whilst Shoemaker takes the ambivalence most people have in their reactive attitudes towards marginal agents to evidence the separation of the three faces, I argue that at least answerability and accountability are unavoidably continuous.⁸⁶

§6.3 Answerability First

I content that responsibility consists chiefly in answerability, and that evaluations of our individual judgements are continuous with evaluations of the regard we hold others in. To hold a person answerable is to demand they tell us how they dealt with a relevant set of instead-of reasons. To hold a

⁸⁵ I take this as one way of failing however Shoemaker distinguishes it as two: we might not notice the facts as normatively salient, or we might judge that whilst these facts initially *seem* like reasons to be quiet, they are merely putative reasons. I take the two together because I take our reasons-tracking systems to be all of a piece for the purposes of responsibility. As we saw in §1.2, sentiments and judgements can check each other, and whilst we might judge that our moral sentiments are misleading us (marking the fact that someone is sleeping as salient, and valencing it in particular ways), this does not mean that we will or ought to be eventually motivated by our judgement. Recall that Huck Finn invites praise because his judgement was checked by his affection for Jim.

⁸⁶ I omit further discussion of attributability and the evaluation of moral character as, contra Shoemaker, I do not believe there are deep selves available for evaluation, and take moral character to only have second-order relevance in responsibility. Whilst further discussion on this point would contribute to a more complete theory of responsibility, it will not specifically advance our understanding of autonomy and the recognition of agential capacities, and so we do not have space to broach the topic.

person accountable is to demand to know how they weighed the instead-ofs grounded in our moral standing or perspective. This provides an overlap but not a continuation. In this section I will expand our understanding of regard by discussing one further way in which we can hold others in improper regard. I will then argue that holding someone answerable involves setting a standard of relevance for instead-of reasons, which trades in precisely the same considerations that regard-evaluating accountability does. Questions of regard are incorporated when we understand answerability this way, and so the two faces of responsibility are continuous.

§6.3.1 Improper Regard

Shoemaker takes regard to pick out the degree to which we recognise, as normatively important or weighty, certain facts about other people, including their perspective on value. Our regard might be impaired because we fail to recognise certain facts about someone as a reason (such as when the noisy passenger didn't notice either that others are trying to sleep or that this was relevant to what his behaviour) or because we mis-weigh that reason (he takes others' sleeping as a relevant but insufficient reason to be quiet). There is, however, another way we can fail to hold someone in proper regard. If we recognise them both as being the source of reasons and being the source of reasons that weigh heavily, we might yet be taking them as the source of the wrong reasons. We might care about this person's interests and perspective *de dicto*, but not *de re*.

Agents can recognise each other as sources of weighty reasons but be mistaken as to which weighty reasons they are sources of. We can care about someone's interests but be wrong about what their interests are. Indeed, the error of assuming that one's own standards of wellbeing, interest, flourishing, or health are what apply to another is what much of the neurodiversity movement rebels against, (Chapman & Carel 2022, discuss an instructive example). Various sorts of oppression occur when people in positions of power, motivated by superficial benevolence, seek to help others live up to the standards they hope to meet themselves.⁸⁷ Concern for others' interests and perspective *de dicto* does not always

⁸⁷ At one extreme, this is the approach to benevolence that is used to justify oppression and colonisation. Measures to 'civilise' colonised people can be motivated by *agape*, but *agape* of a particularly narcissistic species. This love of humanity is love of those characteristics the humans enacting oppression are most fond of in themselves. For a fuller discussion, see Ahmed (2004, 122-143).

lead us to respond to reasons grounded in their interests and perspective *de re*. This is not a failure in judgement, as judgement refers to how an agent handles reasons once they have detected them. This is a failure to recognise which are the normatively relevant facts in this context. Perhaps a neater way to put this is that we recognise someone as being morally important, but under a description that fails to emphasise the right things. Other passengers are sources of reasons, but just not of reasons to be quiet. Intimate relationships are constituted, largely, by certain kinds of attitude, and if these attitudes just don't connect to the person who stands on the other side of that relationship, something has gone amiss and the relationship is impaired. This is, I contend, a kind of improper regard.

In the background of this argument is a broadly Scanlonian understanding of the role of responsibility as serving the maintenance and repair of interpersonal relationships (2008). Relationships with our friends and family are key sites of our flourishing, not only as agents but as complex social animals of the kind we are, and responsibility practices ensure that these relationships continue to function as sites of our flourishing. Scanlon takes particular kinds of regard to be constitutive of these relationships (129) and so holding people responsible restores the relationship to its role by mending regard.⁸⁸ An appropriate description, therefore, is one under which our taking each other means that our relationship can serve its role as a site of our mutual flourishing. When I imagine you inappropriately, this limits our relationship's ability to be that site, and so is impaired regard.

Regard specifies the description under which someone takes us to be on the other side of the relationship with them. When both parties mis-imagine each other, their relationship is impaired. Either can rightly ask whether 'you are answering to me, or to some person you've made up in my place?' To have adequate regard for someone is to take them under an appropriate description. That we are all moral patients and independent sources of moral reasons must feature as part of any appropriate description. That we have an independent perspective on value and that this perspective is owed consideration on its own terms (though is not owed deference) must also feature. The failings Shoemaker highlights pick out certain features of appropriate description against which we might fail, but not all features against which we might fail. The description needs to capture the *contents* of our

⁸⁸ Scanlon allows that we stand in the moral relationship with perfect strangers, however I will limit my remarks to relationships with our intimates: friends, family, colleagues, etc, so as to offer clearer test cases.

intimates' perspectives with some degree of accuracy, not only the status of that perspective as normatively salient.

§6.3.2 Answerability Incorporates Regard

In any instance where one person holds another responsible for some action, the person holding responsible will have an act-description in mind. Just as the act-descriptions we use for ourselves pick out the reasons we see as speaking in favour of some action (Anscombe 2000, §23), the act-descriptions used by people who are holding us responsible pick out the reasons we should not have done whatever we did. Their act-descriptions pick out a set of *instead-of* reasons. The question we are concerned with in this chapter with is how we should establish the appropriate act-descriptions to deploy when considering the agency of a manic person. This is a question of how we appropriately recognise manic agency on its own terms, conferring autonomy, but is also a matter of justice. Whilst we can wrong people by taking a disempowering pathologizing stance towards them in the wrong contexts, we can also wrong them by taking an agentive stance but picking unfair or unreasonable act-descriptions. Bierria outlines the ways in which the sorts of act-descriptions that attach to us due to our social identities can express systemic injustices: in the aftermath of a natural disaster light-skinned people might 'find' food in a flooded shop, whereas dark skinned people are more likely to be described as having 'looted' that food, (2014, 2020, see also Pettigrove 2007, Radoilska 2021a). The social identity 'mad person' is more usually pathologizing, but it can also be oppressive in much the same way. Due to the presumption that mad people did not or cannot handle their reasons correctly, manic people are confronted by less favourable act-descriptions. The terrain is slanted because these act-descriptions highlight sets of instead-of reasons that are unfair, perhaps not even that relevant, or responding to which involves making some other unfair concession. Which 'instead of' reasons are relevant to some judgement is not always transparent. The standard for appropriateness was set, in part, by the context and kind of relationship at stake, and so we will find not only that different interlocutors will accept different reasons, but that different sets of instead-ofs are relevant with different interlocutors.

Consider the case of Sorley McLean, whose letters and poetry written during the 1930s reveal his deep anxieties about whether he should leave home and join an International Brigade to fight in the Spanish Civil war, and what refusal to do so might say about him. Let us imagine he had joined an International

Brigade and had occasion to answer for his decision to a number of different people: his parents, his literary colleagues, and the mysterious *Eimhir Ban* (the young woman with whom he was smitten for much of this time). His parents might readily demand to know why he was risking his life *instead of* staying in safety where he could flourish as a teacher and poet. His literary colleagues might demand to know why he was risking his life *instead of* staying in safety where he would be able to write and contribute to the (fragile, at this time) Gaelic Literary Renaissance, in which he became a foundational figure. McLean indicates that *Eimhir Ban* held some fascist sympathies, so she might have demanded to know why he went to fight *instead of* letting Franco get on with it, or even why he went to fight for the Socialists and Republicans *instead of* the fascists.⁸⁹ In each case, the relevant set of instead-of reasons are established by what the interlocutors care about.

Recall that impairment of a relationship occurs when parties mis-imagine each other, inviting the question 'is it *me* on the other side of this relationship, or some imagined person?' On Shoemaker's account we are answerable insofar as we can report how we handled the relevant set of instead-of reasons. The set of instead-ofs which are relevant is determined in part both by the interlocutor and the kind of relationship, and so the successful demand for answerability turns on both parties being able to agree, perhaps after negotiation, which are the relevant set of instead-ofs.⁹⁰ In order to hold each other answerable, both parties must have adequately good regard for the other, taking them to appear on the other side of this relationship under a description that picks out the things they care about (and whichever other facts ground *relevance* for instead-ofs) with sufficient accuracy. If McLean had been genuinely unable to see why his parents' loss of a child, or the Gaelic literary culture's loss of his contribution, spoke against his risking his life, he would not be answerable on Shoemaker's theory.⁹¹ In

⁸⁹ It seems that McLean held *Eimhir Ban* in somewhat impaired regard. Scholars of his work are confident they have identified the real *Eimhir Ban* and are doubtful that she was in fact a fascist sympathiser, (Whyte 2002, 12)

⁹⁰ Strawson, in the passage quoted above (2008, 9-10, in §6.2) notes that negotiation is available to those who take the objective stance. Perhaps this is true for the kinds of negotiations two small businesspeople engage in when trying to resolve a contract dispute before it reaches court. However, we often engage in other kinds of negotiations while taking the participant stance. The ways we fall into rhythms with our friends and partners, deciding when is boardgame night, who deals with the shopping and who deals with the electricity bills, who picks the kids up from school on which days, are all negotiated outcomes, but negotiations in the sense of being co-stipulated after some give-and-take and trial-and-error.

⁹¹ Plausibly, we should not be able to see why the fascist sympathiser's instead-ofs are relevant. We could not, then, be held answerable by fascists, at least in domains where their fascism influences their instead-ofs. This seems correct to me. That we are immune to being held answerable for our opposition to evil by people who are

order to be answerable, he must be able to recognise his interlocutors not only as the sources of *some* instead-of reasons, but of the particular instead-of that they put forward. Answerability turns on there being a relevant set of instead-ofs, and holding someone answerable can only be done if sufficient regard exists for agents to come to consensus on what instead-ofs are relevant.

Shoemaker anticipates something like my argument (2015, 98). He grants that various failings of regard involve failure of judgement as well, however emphasises that there are distinct objects that are being evaluated here: the agent's capacities to handle reasons, and their capacities to recognise reasons. I do not argue that there are no circumstances under which these objects can be usefully distinguished; I only argue that they are continuous. The distinction, therefore, is only partial. Recall that mania restructures agency not only by changing the normative outlook an agent occupies, (the set of reasons that one has the eyes to see), but also by changing the kind of handling these reasons seem to call out for. Pervasive urgency presents reasons as strict and decisive, and so manic agents handle these urgent reasons as if they had the normativity of requirements. Proper regard marks another not only as the source of weighty reasons, but the source of *these* weighty reasons, which invite handling in *this or that* way. Adequate judgement takes reasons as they are presented and responds to them in the way that they call out for. Impaired regard, which presents one's friend under an inappropriate normative description, is impaired insofar as the reasons one's friend is a source of appear to call out for the wrong kind of handling. These might be reasons to appreciate rather than act, to respect rather than to support, or to relate rather than to listen. These are all ways we can fail to be excellent friends because we are dealing with reasons in the wrong way, but there is not a hard distinction between whether it is our judgement or the normative description under which we take our friend that makes this difference. Judgement and regard are, in this way, continuous.

§6.3.3 Manic Answerability

To be held answerable, as we have seen, involves some version of being called on to give an answer to the Anscombian *why* question. Hieronymi (2013) takes this exchange to be sufficient to be answerable, as having an answer to offer here reveals that "the person has, in some sense, settled for him or herself

trying to do evil is another way of saying that doers of evil can lack standing to hold us answerable for our trying to stop them, (but compare Bailey 2021).

(positively) the question of whether to ϕ ." (14). Such a conclusion reveals the agent's point of view on matters of value, and is what renders them susceptible to moral evaluation. This view of answerability sets a low bar, as any intelligible answer is sufficient to answer the *why* question. Almost any reasons we might offer will be sufficient to count as an answer. Other conceptions of answerability require more robust answers. Shoemaker's account required not only that we reveal what consideration we took to speak in favour of the action, but why we took that consideration to be weightier than some other considerations, in the relevant instead-of reasons. The standard of answerability is higher, requiring that we have something to say about some specific reasons, if our answer is to count. We have discussed Westlund's notion of self-answerability in previous chapters (2009, see §1.4.2 above), but to briefly recap; self-answerability is the disposition to offer a justification for how one sees the normative landscape. An adequate answer does not only include some reason, or refer to some specific set of reasons, it speaks to the adequacy of one's normative outlook as a whole. To be self-answerable is to check in on Tappolet's 'red light', which indicates that we need to exercise regulative guidance. This is a yet higher bar to meet. To hold that manic agents are typically answerable requires that manic agents are not, in virtue of the restructuring of agency we examined in Chapter 2, unable to provide a sufficiently robust answer. Before we consider manic agents' capacities to answer, let us consider these competing standards of robustness.

Westlund's self-answerability is significantly more demanding than either conception of answerability, but this is because Westlund is aiming at another target. Hieronymi and Shoemaker are concerned with moral responsibility, whereas Westlund is concerned with avoiding the particular threat to autonomy that is passivity in the face of one's values. Conveniently, satisfying Westlund's self-answerability is sufficient to satisfy Shoemaker's answerability. Whether the reasons one took to act seemed weightier than the instead-ofs is a matter of how reasons appear to us. Speaking in defence of that overall view includes having something to say as to why these reasons seemed weightier than those instead-of reasons. Similarly, satisfying Shoemaker's answerability is sufficient to satisfy Hieronymi's. By the time we have something to say about why the reason for which we acted is adequate to outweigh certain instead-ofs, we have offered a reason that satisfies the Anscombian *why* question. I join Shoemaker in taking the middle path for the following reason. We are answerable, for the purposes of responsibility, when we can speak to the instead-of reasons, because I take responsibility to be about maintaining and repairing interpersonal relationships. Any reason I could offer might be enough to show that what I did

was intentional (satisfying Hieronymi's question), but whether this intention has any bearing on my relationships with other people is determined by how that reason weighs up against the reasons created by other people's interests, and the value of my relationships with them. Thus, if we are to have our sight fixed on the state of our relationship, we need to be considering various facts about that relationship. I must have something to say about specific reasons if my ability to answer is to play a role in responsibility practices which are restorative. Westlund's standard, however, is too high. Whilst having something to say in defence of one's overall normative outlook is important for one's autonomy, it is too high a burden to set for interpersonal responsibility. Sometimes we do things non-autonomously, but which impair our relationships with other people. In these cases, if we are to repair our relationships, we must be held answerable. The self-answerability standard is therefore too high for present purposes.

This said, I have argued that mania will frequently involve self-answerability. Indeed, it is with attention to the way in which I take self-answerability to function in mania that I motivate the negotiations that are crucial to mad responsibility on my view. In Chapter 2 I argued that mania restructures agency by imposing a pervasive sense of urgency that often causes agents to take reasons as if they were requirements. We considered Jamison's account of one of her own episodes:

"I couldn't worry about money if I tried. So I don't. The money will come from somewhere; I am entitled; God will provide"
(Jamison 1995, 101).

Jamieson volunteers the recognition of her financial situation and specifies that even in the moment that she was aware of it, and apparently of its normative valence (it spoke against her spending), it could not move her in the way she was moved by the need to own beautiful things or set up excellent visual puns. Her financial constraints were a highly relevant instead-of reason, to not spend that money *instead of* spending it. Jamison's articulation here might initially seem to dodge the question, but when we consider the restructuring of agency and the pervasive sense of urgency it involves, the answer is apparent. Concerns about money are normatively relevant, but are pre-empted by the aesthetic considerations in play. Whilst manic agents might struggle to articulate this sense, the answer is implicit in the language of 'demandingness' that pervades the autobiographical literature.

Moreover, is the fact that manic agents seem to frequently be self-answerable. In previous chapters we explored the ways in which these mood states leave intact the disposition to offer justification for how one sees the normative landscape. Westlund notes that this disposition needn't be expressed by answering to people in our social environment - interlocutors in our moral imagination are sufficient, (2009, 36). Thinking about what we would say to some possible interlocutor is sufficient to take ownership of our normative outlook. One explanation of what is happening in manic cases is that the manic agent pays lip-service to the considerations that their interlocutors put forward, but actually takes a different set of instead-of reasons to be relevant. They answer to figures in their moral imagination (perhaps Sappho, Ossian, and Byron) rather than to friends, family, or colleagues in their social environment. They exhibit self-answerability in being disposed to justify their decisions to some imagined person, who will ask why they are wasting an opportunity to write their great novel. Having something to say in defence of one's view of the normative landscape involves having something to say as to why *these* are good reasons, but also why they are better than some particular instead-of reasons. This manic self-answerability, then, satisfies the conditions for answerability in Shoemaker's sense, relative to a particular set of instead-ofs. The self-answerable manic agent can be called upon to justify their action with reference to certain instead-of reasons.

A less schematic version of the same story is as follows: the manic agent imagines their actual interlocutors to share their view of which are the relevant instead-of reasons. This is not unique to mania - we all, mostly, assume our view of what reasons are relevant to be what is shared until we run into disagreement - but the distance between what we imagine and what our friends think is probably much wider. This better explains answerability involving manic people. Manic answerability is directed towards actual interlocutors *imagined* in this somewhat unfitting way, rather than towards Homer or Mother Julian. The frustration manic people experience at not being understood when they practice self-answerability attests to this. Talking to non-manic people is maddening because they seem slow and sluggish and they obstinately refuse to keep up. From a manic perspective, friends and family are being inexplicably mundane about a project that is clearly of mystical or cosmic significance. The intended interlocutor is one's actual peers, but imagined to share one's own ('obvious', but contested) view of what is and is not a relevant consideration.

The instead-ofs the manic person has in mind will not always be the relevant instead-of reasons as Shoemaker intends, which are determined by context, relationship, and other factors (76 f16). However, we can and should recognise that the kind of agency someone has will count among these other factors. The restructuring of agency occasioned by mania alters what is an appropriate description of the person, both as they appear qua agent in autonomy-conferring recognition, and as they appear qua person on their side of interpersonal relationships. The relationship and the kind of agency the agent enjoys both alter what can reasonably be accepted as a relevant instead-of. This does not mean that the manic sense of which instead-ofs are relevant is automatically authoritative, but only that the kind of restructuring occasioned by mania will make a difference for the relevance of instead-of reasons. Likewise, dismissing the manic person's view of which are the relevant instead-ofs is to presume certain neurotypical views of reasons to be the default, which is specifically what we cannot do if we are to confer autonomy on people with BoRD by recognising manic agency for what it is.

So, manic agents typically have an answer to offer, at least under their own act-descriptions, but might these act-descriptions bring with them a set of instead-ofs too far removed from the relevant set to be useful? Might the confidence that mania brings prevent the negotiation I have suggested getting off the ground? In some cases, perhaps. The manic agent might have greater difficulty in recognising and imagining their interlocutor's alternative view. Jamieson could not imagine the reasons not to spend money winning out. Consider, again, Ratcliffe's account of how some intentional states become inaccessible to us, (2010).⁹² If mania is underpinned by some existential feelings which foreclose certain intentional states for us, it might simply be the case that manic agents cannot recognise anything but their preferred sets of instead-of reasons, and where these sets are not the relevant ones, the agent cannot be held answerable. There are perhaps some instances where mania has this effect, so much as to arrest answerability, but we should not understand it to be the typical case. Agents who lose the capacity to hope for things do so, on Ratcliffe's view, because their ability to engage with possible futures has changed (615). To lose hope that one will recover, to use his example, involves being able to assent to the claim that some person with some property might come to no longer have that property, but be unable to recognise oneself as a candidate for being that person. Existential feelings confine our intentional states by altering our sense of the possible, (2020, 251). This is quite different from the

⁹² Again, this is intentionality as *aboutness* (having intentional content applied to an intentional object) rather than rather than to do with intentional agency.

restructuring of agency at work during mania, as put forward in Chapter 2. Manic agents tend to take reasons as if they had the normativity of requirements. This does not make one unable to recognise the normative valence of instead-of reasons, but only to be already committed to a view on why one's own reasons overcome those instead-ofs.

The manic tendency to see the objects of attention as infinitely connected might even facilitate better appreciation of some instead-ofs, if they can be integrated into the grander narratives that the manic sense prefers. Whilst taking the bins out might not feel cosmically important, it is an important part of sharing a life with someone else. If the manic person's love for their partner is of such cosmic importance that it belongs alongside the stories about *Hou Yi and Chang'e* or *Achilles and Patroclus*, the manic agent might be able to render bin-taking cosmic by exercising the *thick fluency* discussed in §3.5. Importantly, relationships that do not involve restructured agency will often still involve disagreements over what is the appropriate act-description or what are relevant instead-of reasons. Neurotypical agents might start closer together, but still face the problem of recognising what kinds of considerations count as relevant in their relationship. They will encounter obstacles when they hold each other in impaired regard, such as when they take each other under descriptions that are too far from what is appropriate.

§6.3.4 The Story So Far

I have argued that whilst the restructuring of agency involved in mania will have a drastic and often detrimental effect on an agent's behaviour, both with respect to their personal interests and their interpersonal relationships, the manic agent continues to be reasons-responsive and so their exercise of these agential capacities demands recognition. Proper recognition is sensitive to the kinds of agency that are in play, and so when agents show up in responsibility attributions under descriptions that do not reflect their particular kind of agency, their agency is not recognised. If holding a manic agent responsible is to confer recognition on their capacities, their agency must be properly represented. Act-descriptions which encode sets of reasons that do not call out for the kind of response the manic agent has to offer risk failing to recognise the manic agent's agency. Settling on appropriate act-descriptions is a difficult part of the usual business of holding agents responsible, and is independently worthy of our attention.

I have motivated a view of responsibility as answerability, where answerability evaluates the quality of an agent's judgement with respect to relevant instead-of reasons, where this evaluation has something to say both about the handling of reasons, which reasons were handled, and what kinds of response these reasons seemed to call out for. Sometimes we evaluate other people's judgement to be bad because they did not properly handle the relevant instead-of reasons, and sometimes because we think another set of instead-ofs were the relevant ones. Either one of these might be the case in mania, as a result of the restructuring of agency that the affective shift produces. Pervasive urgency might cause us to take some reason as if it were a requirement, thus mishandling relevant instead-ofs, or the leap to the cosmic so common in manic thinking might lead us to assume that everyone obviously agrees with our view of the relevant instead-ofs. However, the non-manic interlocutor who holds a manic person responsible does not have unilateral authority to decide on act-descriptions. We then face our final question: how, then, should a manic and non-manic person negotiate the relevant act-description for the purposes of answerability? Answering this question gives us the final piece we need to inform a view of what proper recognition of agential capacities looks like in the context of autonomy.

§6.4 Fixing Act Descriptions

What is the point of holding someone responsible? In §6.3.1 I noted that I take a broadly Scanlonian view of the role of responsibility: holding each other responsible is how we maintain, transform, and repair our interpersonal relationships to ensure that they continue as sites of our mutual flourishing. Responsibility practices, such as remonstrance or apology, are appropriate insofar as they contribute to this repairing or transformational work. Repair restores a relationship to a well-functioning example of its kind, and transformation changes the relationship to a kind that it might better function as. When we hold people responsible, we often contest act-descriptions. Consider the ubiquitous example of the lovers' tiff, where one party (for now, Alice) demands to know why their partner (Betty) was flirting with someone else, and Betty denies that she was flirting. The act-description is important to these characters as how the act is described has significant consequences for their relationship. Thus, if they are to maintain their relationship in an unimpaired state, they must find a description under which to

recognise that act, and possibly alter features of their relationship to accommodate the act under that description. Perhaps they will agree that Betty was flirting. If they have been maintaining a closed relationship so far, they might agree that this is a betrayal of certain principles of that relationship, and that amends should be made. Perhaps, after some discussion, they will instead decide to maintain an open relationship. Alternatively, they might decide that this is an unavoidable part of Betty's nature, accepting that Betty will flirt with people sometimes but that it has no implications for their relationship. On the other hand, they might decide that Betty was not flirting and recognise that Alice has certain insecurities that are triggered by innocuous interactions. Either of them might suggest that they take steps to help Alice in overcoming these insecurities and help Betty to avoid triggering them. On yet another hand, they might discover that Alice has some slightly possessive and controlling tendencies, and that Betty needs to establish clear boundaries if the relationship is to continue being healthy. Depending on the act-description that is selected, different solutions present themselves. In each case, the objective is maintaining the relationship. Even if the partners break up, they are transforming their intimate romantic relationship into something more like the relationship between moral strangers, but there is still a basic relationship there that can be a site of their mutual flourishing.⁹³ If they fail to find an act description, if Alice doesn't broach the subject and instead pretends to herself that Betty is not an incorrigible flirt (she's just friendly, and everyone else has a misjudged sense of boundaries, perhaps) then this is likely to impair their relationship. It will be a less effective site of her and Betty's mutual flourishing.

This is the basic model of virtuous holding responsible that I have in mind. Alice and Betty are likely to have a fairly straightforward task in negotiating act-description as both share the concept of 'flirting' and, whilst they might not share each other's view of the boundaries of flirting nor the role it plays in romantic relationships, they can easily become familiar with the other's view of these boundaries and roles. The matter is complicated in the case of mania by the cosmic fixation, the thick fluency, and the sense of pervasive urgency that manic people bring to agency, and by the presumption of unintelligibility or illegitimacy that is attached to mad perspectives. Two features are necessary for responsibility ascriptions to confer recognition, on my view. The first is that parties approach the matter

⁹³ I do not have space to defend this position here, but I take our relationships with moral strangers to be sites of our flourishing in the minimal sense that they are sites for us to exercise virtues by recognising the independent value of other people and their projects. This is, in some ways, analogous to the *civic friendship* that holds between members of the same polity who have never met, (Aristotle 2011, 1242a).

with something like the attitude that Rashed calls *reconciliation*, (2019, 114). The second is to jointly decide how the act is to be deployed in the narrative of the intimate relationship, stipulating the description that parties to the relationship apply to the act.

This decision is guided by the norm of mutual flourishing. The norm means that harm that is done cannot be ignored, but it sometimes means that amends needn't be made. As Walker puts it:

Sometimes we pardon or excuse, deciding that the offense does not require redress, or that it is better to forgo redressing it, or that although redress is in order, there are reasons to be merciful and demand less than is owed.

(Walker 2007, 10)

I take typical intimate relationships to include a commitment to sometimes take on more than is one's share of emotional labour.⁹⁴ Sometimes we will be sad, or tired, or beset by futility thinking. Being unable to see the value in our friendships, we will not do our part in keeping them up. This is where our friends step up to carry more than their share. As we will all fall short in our friendships sometimes, we will all likely benefit from this norm, at some point in our life, as well as carry the burden for it. Whilst friendships should avoid a situation where one person is always doing significantly more work, I take friendships to usually involve this commitment to not pay too much attention to who is doing the hard carry.

§6.4.1 Reconciliation, Understanding, Empathy

In order to fix act-descriptions in a way that facilitates the reparative work of responsibility, agents must approach act-description negotiations with the right attitude. The attitude required to facilitate this negotiation is something like the attitude Rashed calls *reconciliation*. It is an attitude which grants that one's interlocutor's perspective, their act-description or preferred set of instead-of reasons, is at least a legitimate starting position in negotiation, but one which retains space to reject that perspective in the

⁹⁴ For clarity's sake, intimate relationships are distinguished from the basic moral relationships that holds between strangers. Platonic friends, lovers, cousins one sees once every few years, and casual acquaintances all count as intimates for our purposes, not just romantic partners.

final evaluation. As I explore below, parsing exactly what this attitude involves is tricky, as it must be carefully distinguished from some near neighbours.

Rashed describes reconciliation as both a desirable state of affairs reached through a particular process, and an attitude that is required if that process is to be effective. To hold this attitude is “to approach others with the implicit acknowledgement that their satisfaction counts, that their ability to find themselves at home in the world matters as much as mine does” (2019, 116). Both ‘satisfaction’ and being ‘at home’ involve others’ acceptance and acknowledgement of one’s identity and projects, and of the worth of each. The world here is the shared social world, which can be distressing, disempowering, and sometimes dangerous for us if our identities do not fit. Reconciliation as process is part of the wide social project that mad activists are engaged in, and amounts to an exercise in hermeneutic power – the power to create social meaning. For mad activists to be received with an attitude of reconciliation is for them (and the mad social identity they represent) to enjoy both the power to create social meaning, and the standing to contest how they exist in the social imaginary.⁹⁵ As Rashed has it, mad people must be able to write mad narratives into the shared cultural repertoire in order to be at home in the social world. Reconciliation is an attitude that participants, in a cultural negotiation as to what social categories like ‘madness’ mean, must hold. What, then, must be the case if the private counterpart to this public attitude is to obtain?

Rashed suggests that it is not typically appropriate to ask why we should approach our friends with an attitude of reconciliation, such an attitude is assumed, (117). This recalls Scanlon’s view of intimate relationships as being partly constituted by particular attitudes, (2008, 131). Certainly, we are usually open to learning more about how our friends see the world, and two people who are unwilling to offer the kind of recognition that this attitude makes space for are not really friends. Despite this, we do not always occurrently hold this attitude, and holding it does not always secure recognition. We can be

⁹⁵ This is a distribution of hermeneutic power. Coady (2017) considers maldistribution of hermeneutic power to be a distinctive kind of epistemic injustice, and is one that is widely endured by mad people. Insofar as the social identity ‘mad person’ is understood in terms of incompetence, it is a social identity that enjoys less hermeneutic power than is due. The activism Rashed examines is an exercise in using what limited hermeneutic power mad people have in order to change the content of that social identity so that it will be less subordinating. This is an uphill struggle, as changing the content of a social identity requires precisely the power that mad people are deprived of by this maldistribution.

deeply invested in our friends' satisfaction and at-homeness in the social world, whilst failing to make space for, or acknowledge as valuable, their actual identity. The manic person might see their friends' satisfaction as deeply valuable yet act as if their friends' actual projects were trivial, assuming that the manic agent's own (urgent and obvious) sense of what is valuable is shared. The non-manic person might be deeply invested in recognising their friend's identity, but only their identity as it would be articulated while not manic. More than investment in our friend's satisfaction and fit within the world *de dicto*, we need an appreciation of what, *de re*, their satisfaction or belonging in the world would consist in. Reconciliation is a corrective attitude that is actively deployed, rather than being a constant background condition. Good friends have this attitude available as a reparative resource, to be activated in contexts where they are not 'meeting' each other properly in the relationship. Such a failure to 'meet' is an impairment, but unlike the backbiting friend from Scanlon's example (2008, 129), it is not a kind of impairment that makes us question whether the two are really friends.⁹⁶ The attitude is available and plays an important role in friendships but is not constitutive of the relationship in the way that attitudes of fondness, trust, and loyal hope are.

Might something more familiar than reconciliation do the crucial work of facilitating this negotiation? Olivia Bailey has argued (2021, 2022) that empathy is a crucial element of *humane understanding*, which is itself important for our flourishing. Might humane understanding, or the empathy that underpins it, do the work of facilitating these negotiations? Humane understanding is a "direct apprehension of the intelligibility of others' emotions" (2022, 51). She argues (and I am sympathetic) that humane understanding is a non-instrumental moral good, but we might also consider it an instrumental moral good if it facilitates virtuous responsibility negotiations. I am reticent to accept this move, however. Bailey understands empathy as involving endorsement of the feelings that are empathised with: it "typically entails emotionally apprehending the object of their concern (or, rather, the imaginative analogue of that object) in the same evaluative light that they do." (2021, 9631). Someone's empathy with our emotional experience "secure[s] a first-hand appreciation of the apparent appropriateness of that emotion" (9638). Humane understanding depends on empathy, (2022, 57) and empathy for the manic agent, therefore, involves granting that the manic view of reasons is *appropriate*. Virtuous

⁹⁶ Schulman describes something similar when she looks for 'agreement' in conflict: "Not that we would hold the same views, but rather that we would communicate enough to agree on what each of our different views actually are. That I could tell you what you think and you would recognize my rendition as accurate, and you could do the same for me" (2016, 150).

negotiations require that the manic sense of things be admitted as a legitimate starting point, but it also requires that the possibility of rejecting that perspective in the end. Empathy understood as endorsement, therefore, is too strong to permit a virtuous responsibility negotiation. It might be suggested that the agent holding their manic friend responsible might be able to endorse the manic perspective without giving up their own opening position – endorsing both views of reasons as appropriate (relative to different identities or sets of projects, perhaps). This empathetic agent can take a leading role in this negotiation, perhaps guiding their friend to a more virtuous perspective. This might be achievable, however for the purposes of responsibility we must acknowledge that sometimes others' perspectives will not represent reasons appropriately. Empathy as endorsement is not a suitable attitude with which to approach responsibility negotiations as it does not leave the proper space for us to eventually reject the perspectives of other agents.

If we continue to look for more familiar notions than reconciliation, we might opt for understanding, or we might hold onto empathy but reject the claim that empathy requires endorsement.⁹⁷ Schramme (ms) makes both moves, focusing on understanding as it appears in questions of hermeneutic justice. (Reconciliation was put forward as a shared social attitude for the purposes of exercising a power to create social meaning, so private interpersonal hermeneutic justice might be a good place to look for a private counterpart). Schramme argues that the kind of understanding we are looking for when we try to 'get the point' has a crucial normative element: "To understand another person... it is often required to see their perspective as valid in a specific way." (7) Might this specific way of seeing a perspective as valid mean that we accept it as a valid starting point for the purposes of negotiation? Schramme tells us that:

For understanding to fail permanently, a perspective must be such that it cannot be imagined to be a humane perspective; it must lack a connection to a normative account of what it means to be a normal human being.

(Schramme, manuscript, 9)

⁹⁷ If humane understanding is still a contender, it is because we have removed the requirement that it be underpinned by empathy as endorsement, and lean more on its representation of others' perspectives as intelligible. Schramme's account of understanding has this structure and so is the better model with which to examine this move.

He offers the example of Mr F who, when shopping for clothes, required to stand by large windows in shops in order to commune with the goblins in his head, who gave him advice on such matters. Mr F's perspective is understandable, Schramme points out, and indicates that "we do not need to consider other people or their mental states as rational to minimally understand their perspective." (10). The attitude of understanding is one that anyone should be able to adopt towards Mr F. Might this offer a better facilitating attitude for responsibility negotiations? I have two concerns. The first is that Mr F's perspective was not a perspective on reasons. His goblins advised him on what trousers were suitable, but this is perfectly understandable either as an instance of taking advice (a commonplace activity, even if we must grant the goblins some independence of Mr F), or as an instance of deliberation (Mr F is reflecting on the relative weight of reasons, using the goblins as mouthpieces in his internal dialogue).⁹⁸ What is less obvious is that this kind of minimal understanding would attach to Mr F if it was his perspective on what counts as a good reason that was esoteric, rather than his means for handling reasons. If Mr F considered, while shopping for clothes, how well each piece of clothing would function as armour to protect his vulnerable chakras, he might not be understood (especially if his description of chakras did not track familiar religious ideas, or he did not articulate the dangers or protective properties of clothes in ways his interlocutor had the concepts to follow). Schramme notes that, if Mr F had insisted that he must stand on his head before buying clothes, he would be less likely to be understood. Mr F's goblins are understandable as features of a perspective on what was happening. Perspectives on what is a reason are more at risk of non-understanding, and it is these perspectives that matter for responsibility negotiations. This takes us to my second worry, which is that the kind of validity Schramme notes seems to fall below the standard required to facilitate negotiation. This validity is achieved when an interlocutor exerts some effort to see (for instance) Mr F's behaviour as an unusual instance of something familiar. We can, however, align our perspectives in this way with perspectives that are not legitimate opening positions in a negotiation. The validity conferred on a perspective for the purposes of understanding marks it as within reach of some normalcy. A great many illegitimate opening perspectives are within reach of normal ('my interests count morally and no one else's do,

⁹⁸ Taking advice from voices in one's head, whether one considers them to be renegade thoughts or to belong to some other entity, is a popular way of accommodating these experiences. Much in the same way we might ask ourselves what our anger at our friend, who has done nothing wrong, is telling us (perhaps that we are angry at ourselves), people who hear voices sometimes interrogate those voices for what their mind is trying to highlight to them. For brief introductions see Eleanor Longden's Ted Talk (2013) and Rashed (2019, 192). A deeper exploration is available in the documentary *Healing Voices* (Moynihan 2016).

therefore everyone should give of themselves to tend to my needs’).⁹⁹ Holding mad agents responsible is a way we might include abnormal perspectives in restorative moral practices, but we must still be able to filter out morally improper perspectives (whether they be mad or neurotypical). The validity standard here is unsuitable for our purposes.

Empathy as endorsement is too strong to allow for useful negotiations of act-description, and basic understanding is too weak to rule out the genuinely inappropriate starting positions. What we needed was something that grants the legitimacy of the manic perspective as an opening position and retains space to reject that position, in order to achieve repair in a relationship. For this reason, reconciliation remains our best model. It is specifically geared towards a conclusion as to how something (identity in Rashed’s use, act-description in ours) should be understood, guided by a need for everyone to be at home in the shared social world. When adapting this attitude to the private interpersonal context of responsibility negotiations, the shared social world is replaced by the relationship in question. A relationship in which we find acts attributed to us which we do not recognise, or which are incompatible with our self-conception, are relationships in which we are not at home. Sometimes this means we must revise our self-conception to bring it closer to those which are ascribed to us by our friends, but sometimes our friends need to realise that they aren’t seeing us straight. Private reconciliation, then, is an attitude that prioritises our friend’s being at home in our interpersonal relationships. It is a resource that is on hand for intimates, but is not always active. Schramme notes something important here, which is that basic understanding is an achievement. It is sometimes effortful. Some people will not passively understand Mr F’s perspective as valid, but can come to do so by aligning their perspectives for a moment. We need to overcome the tendency to assume that our own view is shared or is obvious. Underlying this attitude of reconciliation is a different kind of empathy, which Zahavi and Fernandez call *basic empathy*. They write:

Empathy is *not* about me having the same mental state, feeling, sensation, or embodied response as another, but rather about me being experientially acquainted with an experience that is *not* my own. Empathy targets foreign experiences *without eliminating their otherness*. In empathy, I am confronted with the presence of an experience that I am not living through myself. If I empathize with your sadness, I

⁹⁹ I leave aside how we should hold responsible someone who offers this opening position. This attitude is, crucially, not consistent with the attitudes that constitute intimate relationships. This person might be very familiar, but as they are taking an objective stance towards us, we are (at best) moral strangers.

have a sense of what it is like for you to be sad without being sad myself; I lack first-person access to the sadness in question. To empathically grasp another's joy is not to be joyful oneself, but to recognize the joy as *belonging to the other*.

(Zahavi & Fernandez 2020, my italics except for the repeated 'not' in the first sentence).

Basic empathy is the recognition of another's perspective as genuinely other, belonging to another, and does not assume endorsement. To empathise with someone in the basic sense is to recognise them as independent, and not a reflex of the self. We often fall short of this basic empathy, such as when we assume others see matters in our terms, filtering their testimony through our own values or concepts. We do this when we assume that our sense of what reasons are relevant is obvious or uncontested. Reconciliation requires that we overcome the tendency to think of our own experiences as expressively adequate, and to look for what is *other* in other people's experiences.

Friends, siblings, lovers, and other intimates sometimes need to activate the attitude of reconciliation, not because they fail to be genuinely close friends, but because basic empathy is a frequently effortful achievement. The acknowledgment that needs to be made for interpersonal reconciliation to function is the recognition that another person's perspective does not need to come in terms that one is familiar with – one's own normative outlook is not expressively adequate for all of normativity. An appropriate attitude of reconciliation requires that one confers on one's interlocutor the *standing* to contest both positions in, and the terms of, a discussion, as well as the *power* to stipulate shared meaning. How might we achieve this, particularly when feelings are running high, and we are not thinking about standing and meaning-making power? Jim Sinclair's entreaty, with which I opened this enquiry, would be an excellent conceit: "[g]rant me the dignity of meeting me on my own terms – recognize that we are equally alien to each other, that my ways of being are not merely damaged versions of yours" (1992, 302). Neurotypical people who are holding their manic intimates responsible must, if they are to maintain their relationship as a site of mutual flourishing, meet their manic friends on their own terms. Yet, manic people also need to recognise that whilst it seems that their non-manic friends are slow, sluggish, and obstinately mundane, that theirs remains a genuinely independent perspective on what matters. Remonstrating with someone is one such method of signalling that basic empathy needs to be activated, however isn't always an effective one. How best to embed this conceit into responsibility practices will vary from relationship to relationship, but where it needs to be actively supported, practices such as calling 'timeout' or marking some shared private reference can communicate that it is

time to adopt the reconciliation attitude. Figuring out how to do this is an important, though often subtle and implicit, part of any relationship. Insofar as a bipolar agent's friends and family play an important part in their self-care regiment, it is commonplace and valuable that red-flag phrases such as this emerge. Where they do not emerge organically, however, they can be stipulated during straightforward and honest communication.

§6.4.2 Deploying an act under a description

By approaching responsibility ascriptions with the right attitude and with basic empathy activated, both parties are able to recognise that the other's act-description is relevant to what kind of answer is called for in holding one another answerable. There are, then, at least two act-descriptions in play, and as a discussion (like that Alice and Betty were having) continues, several more act-descriptions will probably emerge that might lie somewhere between the two starting positions. How should our negotiators rule these candidates in and out? If we take responsibility ascriptions to maintain interpersonal relationships as sites of mutual flourishing, there are three reasons to rule a description out. It might cause someone to show up under an inappropriate description on the other side of the relationship; it might damage the relationship, occasioning excessive recrimination or causing us to give up on each other; and it might cover up problems in a relationship, which might fester over time. In each case, an act-description is inappropriate as it fails to maintain the relationship as that site of mutual flourishing. I will address each in turn.

The first condition under which an act-description can be ruled out is if it is incompatible with one of the parties showing up, under an appropriate description, on their side of the relationship. Act-descriptions which do not accommodate the restructured agency of mania are inappropriate because the relationship is with someone who, as a matter of their existential reality, will experience this manic restructuring of agency from time to time. Thus, the maintenance of the relationship must accommodate that fact. Act-descriptions which are incompatible with this recognition fail insofar as the manic person can reasonably ask 'is it me on the other side of this relationship, or someone you have imagined?' Act-descriptions might fail to recognise this agency if they assume a neurotypical view of which instead-ofs are relevant to be default. A relationship is not an effective site of mutual flourishing if the parties 'miss' each other in this way, either taking the other under an inappropriate description. In

such a relationship, some repair is needed more than maintenance. Negotiating act-descriptions is an opportunity for both parties to refocus the descriptions under which they take their friends, carrying out that repair by mending the regard they hold each other in. Act-descriptions, insofar as they encode sets of reasons which speak in favour of the action and instead-ofs which speak against, can fit more or less well to the appropriate descriptions of the agents to which the acts are attributed.

Act-descriptions are also inappropriate if they endanger the relationship as a site of mutual flourishing by introducing excessive recrimination, or risk causing the parties to give up on the relationship. Insofar as such a relationship is no longer a site at which both parties flourish together, picking this act-description as part of holding responsible makes it harder for responsibility practice to serve its purpose. This is not to say that we should avoid recrimination and should never give up on a friendship – very frequently we should – it only means that the objective in holding someone responsible is governed by a flourishing norm. Sometimes it is better that we become moral strangers, as we are so incompatible or have hurt each other so much that no other sort of relationship is one in which we could mutually flourish. This is a drastic transformation response, however in most cases maintenance and repair allow us to keep our relationships as the same kind that they are whilst ensuring we can flourish together in them.¹⁰⁰ This flourishing is interrupted by an excessive focus on who has fallen short of what is expected of a friendship. As outlined above, friendships and other sorts of intimate relationships must avoid an excessive imbalance in who is doing the work but, I expect, usually involve a commitment to not track too closely how the balance sits. This norm earns its keep by making sure that friendships are resilient enough to weather times when one person isn't showing up for the relationship. At some time or other we will all be that person. In cases of mania, the particular form of not showing up is likely to involve impaired regard (assuming our friends share our cosmic sense of what is important) or in the handling of instead-of reasons (taking them to be lightweight due to the pervasive sense of urgency that mania brings). These are specific failings to do our part in a relationship, but they are neither unique nor more egregious than is commonplace, and the norm of looking at the balance in a relationship at a generously fuzzy resolution can absorb these shortcomings. Act-descriptions such as 'you dismissed me as a consideration' are useful for prompting someone to revise their behaviour and might bring someone to

¹⁰⁰ And even when such a transformation allows the moral stranger relationship to be a site of our flourishing, it is still probably the least bad failure state for the relationship we had previously enjoyed.

revisit self-management practices that are not working, but they can sometimes focus too strongly on the shortcoming to do effective reparative work.

Finally, we sometimes describe the things we do in ways that obscure problems or hide our wrongdoing. Sometimes we accept these descriptions because it seems less troublesome for us than insisting on act-descriptions that capture the normatively important features more accurately. The term ‘toxic positivity’ refers to the tendency to focus on superficially positive descriptions of acts, circumstances, motives, etc, and is often expressed to characterise narratives which lionise gratitude and insist that ‘happiness is a choice’. A relationship beset by this toxic positivity is one in which the parties refuse to acknowledge the problems in their relationship, glossing over them with superficially positive descriptions, or choosing to describe any features of their relationship but those which endanger it. Whilst more favourable act-descriptions can support our flourishing by avoiding unnecessary recrimination, they can also poison the well. Problems other than misrecognition or excessive recrimination cannot be removed by description choices, and so when our relationships are endangered for reasons other than our act-descriptions, the solution cannot be found in better act-descriptions. Pollyannaishness or toxic positivity, endangers our relationship’s status as such a site because it diverts us from the reparative work that needs done. Thus, act-descriptions that fall into these trends must be ruled out as they are obstacles to, rather than means to, maintenance and repair.¹⁰¹

In each of these cases an act-description is inappropriate because maintenance and repair is obstructed by narrating the relevant part of the relationship with that act-description in play. These negotiations are opportunities to decide, for the purposes of this relationship, what was done. The norm of flourishing governs this decision, however it does so loosely. These act-descriptions are ruled out, but the norm does not offer clear guidance on which descriptions to rule in. This is as it should be. We all have wide latitude to craft the specifics of our relationships by deciding how to describe various acts or events. The minute textural differences between two relationships are the product of the choices that a person and their friends have made about how to understand different events, actions, and contexts,

¹⁰¹ Sarah Schulman (2016) describes a similar phenomenon, as a feature of *supremacy*. Supremacist self-conceptions are resistant to change because interpretations of events according to which the supremacist invites criticism are not only rejected, but are treated as transgressions themselves, (59).

and indeed on which events, actions, and contexts are significant enough to demand understanding. Relationships in which one person experiences episodic mania and the other does not will require particular kinds of understanding that relationships between two neurotypical people will not. There is wide latitude for people who experience mania and their friends to work out the specifics of their relationships but if the relationship is to be a site of mutual flourishing, they are confined in the way outlined above. They must synthesise the manic sense of things with the non-manic in order to accommodate both parties' perspectives.

I will briefly note that, whilst it would likely be harmful to a relationship to constantly re-negotiate act-descriptions, negotiations can be rerun under some circumstances without creating a problem. A person who is manic might struggle to remain focused, or to properly activate basic empathy, during the intense revelatory experience of mania. It might be the case that holding responsible has to wait until someone is less swept up in a flight of ideas, before engaging in repair. If some repair must be done immediately, but the negotiation is hampered by an agent's mania, it might be best to engage in an initial negotiation, carry out the initial repair, and then wait until one's friend has a firmer grip of their mania or the episode remits to re-negotiate the act-description and engage in a more thoroughgoing repair. Delayed re-negotiation will also be useful in cases where the parties initially settled on act-descriptions that papered over cracks in their relationship, but which only became apparent to them later. The maintenance of relationships is an ongoing process and sometimes things that need to change include the act-descriptions we had previously settled on for important matters.

§6.4.3 The Moral Relationship

It might be objected that the model I suggest is applicable only in intimate relationships where the attitude of reconciliation with basic empathy can be expected or negotiations could take place. There is no guaranteed space for these features when we hold moral strangers responsible and so act-description negotiations cannot take place. It is not my purpose here to defend this model in the moral relationship, as this would require a detailed account of what the moral relationship involves that I do not have space to develop. What I will note, however, is that the kind of approach I am describing is available to moral strangers, assuming a degree of success in the mad activism that Rashed examined.

The private meanings produced in interpersonal relationships will not be available between moral strangers, however if mad narratives can be written into the shared cultural repertoire, then non-manic strangers will have at least some useful resources available with which to approach responsibility practice with manic people. The kinds of private meaning-creation I outlined above are, in fact, a grassroots version of the production of social meaning that is needed for these narratives to enter the shared cultural repertoire. The local conferral of hermeneutic power is one means by which mad narratives, suitable to inform responsibility practice, might be generated before they are taken up by mad activists and presented to the whole community. Whilst private exercises of meaning-creation will not disseminate narratives very quickly, the production of specifically suitable narratives to include mad people in the reparative exercises of holding responsible are valuable and must be created before they can be written into any cultural repertoire.

§6.5 Constitutively Social Autonomy

This chapter has considered how our responsibility practices confer recognition on agents. If, as we have assumed, autonomy requires the recognition of an agent's reasons-responsive capacities being at work, then it is important that the capacities the agent has, their particular kind of agency, is what is recognised. 'Agent' is a description that we cannot dispense with when interacting with others, but it is an underspecified description and so to confer autonomy by socially recognising someone's agential capacities at work, we need to recognise manic agency when it presents itself, rather than mistake it for a defective kind of neurotypical agency. Responsibility practices are a vehicle by which we can offer the kinds of recognition that is appropriate given the restructuring of agency that mania occasions. Negotiating the description under which an act will be understood, for the purpose of responsibility, can incorporate the manic perspective on reasons, and so extend appropriate recognition to manic agency. Virtuous responsibility practices are those which include mad agents in the repair and restoration. Recognising manic agency on its own terms allows responsibility attributions of individual acts to be local instances of this general autonomy-conferring recognition.

§6.5.1 - Responsibility and Autonomy

Before elaborating on this point, it is worth saying a brief word on the relationship between autonomy and responsibility. In §0.1.2 I outlined the model of autonomy I would use as a vehicle for this investigation, according to which autonomy requires agential capacities to operate, and to enjoy social recognition. The model is useful as it can be adapted, *mutatis mutandis*, to more closely mirror other theories of autonomy which strike different balances between ‘internal’ and ‘external’ features of autonomy. As I have set it up, the relational side of autonomy is a socially conferred status, involving recognition of someone’s agential capacities either generally (globally), or being at work in this (local) instance. On the account of responsibility I have offered here, we must at least confer local autonomy if we are to hold someone responsible, as we must recognise their agential capacities at work, taking an agentive stance towards them rather than a Strawsonian objective stance. Local autonomy is necessary for holding a person responsible. We can, however, hold someone responsible without recognising them as a responsible agent. We do this if we only recognise their agential capacities at work in a few local instances, but do not recognise those capacities as being general features of them as people. Plausibly, we sometimes should do this. Small children are in the process of developing the agential capacities we hope to recognise as autonomous in the future, but too much recognition can be both unjust and harmful. It is unjust as it *overagents* children, taking them under an unfitting description that places unreasonable expectations on them. The practice of trying children as adults, which exists in a number of jurisdictions, is plausibly the key example of this *overagenting* injustice. Being too ready to recognise global agential capacities is also potentially harmful to the development of a child’s agency, as their abilities to track and respond to complexes of reasons can easily become swamped when they are expected to track and respond to many more considerations than they are familiar with.

On the other hand, holding children responsible involves helping them to see the act-description, and implicit set of instead-ofs, that speak against acting as they did. Insofar as this is a nurturing relationship, the objective is to help children learn how to respond to wider sets of (instead-of) reasons. Responsibility’s role in moral education is not to instruct us as to what we should and should not do, but rather is to help us develop the eyes with which to see more kinds of reasons. In the case of adults, whether they are mad or not, the same process is at work, however usually in more localised contexts. When our intimates, or even moral strangers, hold us responsible, they draw our attention to sets of instead-of reasons that they think we didn’t see or did not weigh correctly. When they are correct, we

are being called upon to recalibrate our reasons-responsive capacities.¹⁰² Maintaining relationships with friends involves sensitising ourselves to their wants and needs. This process involves developing modes of agency appropriate to be that person's friend, which is to say, sensitivity to such reasons. Developing these modes and keeping them as part of our inventories of agency is an exercise in cultivating our overall agential capacities. Thus, being held responsible serves to improve the agential capacities that ground autonomy.¹⁰³

Holding an agent responsible does not assume global autonomy. Being held responsible does not require that we have general agential capacities, only local ones. Neither do we need to be recognised for any general agential capacities we do have, if we are being held responsible (locally) for some action. However, if we have these capacities and they go unrecognised, the relationship we are in is in an impaired state, as we appear under an unfitting description. Our friend does not properly 'meet' us in this relationship. Therefore, one of our objectives in negotiating act-descriptions when we are held responsible in this way will be to correct the improper regard in which we are held. Correcting this misrecognition secures our relationship as a better site of flourishing, but it also confers recognition on our general agency, supporting one's global autonomy. Being held responsible confers autonomy locally, but also creates the opportunity to demand that global autonomy be conferred. There is a cyclical and virtuous relationship between autonomy and responsibility as I have presented it. Being held responsible secures minimal, local, autonomy, allows us to sharpen the agential capacities that we have, and pursue greater recognition of those capacities during negotiation. Those who recognise our

¹⁰² The analogy I draw between children and adults stands on the shared objective of sharpening agential capacities in both cases, an objective it shares with some recent work in psychiatric ethics (Pickard 2011, 2017, Brandenburg & Strijbos 2020). In Brandenburg and Strijbos' case, they refer to the appropriate stance as one that is nurturing. Attractive as this sounds, it might be objected that it is unduly infantilising to treat adults in the way children are treated, and worse in the context of mad adults as this activates already threatening tendencies to unagnet mad adults. In response, I emphasise that the project of developing our agential capacities is never complete, and whilst we might flatter ourselves that as adults we only need to develop local modules to suit particular relationships, it is probably not the case that we are as sensitive to reasons as we *should aspire* to be. The virtuous person is a regulative ideal, and so it is not realistically possible that we ever get there. Consequently, our task is to continually work in our own improvement. I am grateful to the *Philosophers with Lived Experience of Mental Disorder network* for fruitful discussions on this point.

¹⁰³ In this sense, being held responsible still contributes to satisfying social conditions for autonomy, but formative rather than constitutive social conditions.

agency are taking an agentic stance towards us, and so are more likely to hold us responsible in future. Thus, the relationship is cyclical and self-reinforcing.

It is highly intuitive that things which we do autonomously are things we are responsible for, and indeed that the more often it is fitting to hold us responsible, the stronger our global autonomy is. On the view I have offered here, this is probably a contingent truth, but not a conceptual truth. The practices of conferring autonomy and holding responsible are mutually re-enforcing, but not co-extensive. We confer autonomy by holding agents responsible because in both cases we must recognise agential capacities at work. Responsibility attribution is not the only possible source of this recognition, but it is sufficient to confer some basic recognition, and so conferrence of (some, local) autonomy is intrinsic to holding responsible, but holding responsible is not intrinsic to recognition of autonomy. Oshana (2002) and Fischer (2012a), who interestingly sit far to either side of the internal-relational spectrum from each other, both resist the intuitive link between autonomy and responsibility. Oshana holds a strictly relational view of autonomy, according to which the degree of autonomy a person enjoys is determined by social-material facts about which ends they can effectively pursue, (2002, 273). On such a view, it is perfectly natural that our capacities to answer for what we have done as part of a moral conversation come apart from our autonomy, (270). A slave, who definitionally lacks autonomy, can still be expected to answer for their behaviour in their (perhaps few) private or undirected hours (277-8). On the model I use, recognition is the basic social condition for autonomy, however it is not exhaustive. We can recognise agential capacities even where social conditions such as poverty or bondage drastically limit the range of responses the agent can bring to bear on reasons. The role of recognition keeps the model of autonomy I am using further to the centre of the internal-relational spectrum, but the model is flexible and can be moved around that spectrum by specifying other social conditions. As long as recognition is a basic social condition for autonomy, however, some intrinsic link is preserved between holding responsible and conferrals of autonomy, but the relation only goes in one direction. At the other end of the spectrum, Fischer holds that autonomy consists in being governed by one's deep or true self, (2012a), and that there is some metaphysical fact of the matter as to whether one is (can be rightly held) responsible for some action, determined by whether there are possible worlds across a given domain at which, holding one's reasons-responsive capacities fixed, one acted otherwise, (2012b, Fischer & Ravizza 1998, 64). He emphasises that the designated part of the mind or self that sets normative standards in autonomy and responsibility cases need not be the same, (2012a). The practical

identity is what is designated for identification in matters of autonomy, whereas the mechanisms that issue in action are designated as the objects for responsibility. The partial overlap between autonomy and responsibility on these views is unsurprising as practical identities and action-issuing mechanisms are both objects that trade in practical reasons, but do so to different ends. Consequently, the overlap is largely incidental.

My view, along with those of Oshana and Fischer, leaves a noticeable gap between responsibility and autonomy at the conceptual level, but allows for points of connection as described. Such an interpretation, however, has received pushback. Radoilska (2021b) has argued that whilst autonomy and responsibility aim at different normative objectives, we should not give up on an important normative connection between them. Whilst these connections are highly variable according to different conceptions of responsibility or autonomy, a detailed analysis of these conceptions show that some connection is reliably retained. The view I have offered is consistent with this picture, with the qualification that the connection is unidirectional. Ascriptions of responsibility recognise agency and so satisfy the minimal social conditions for autonomy, however neither this social recognition, nor recognition of the robust autonomous agency that it permits, is sufficient to mean that an agent is being held responsible. Autonomy and responsibility, on this view, are continuous going in one direction, but not in reverse.

§6.5.2 - All Tolled

In Chapter 2, we established that mania restructures agency by altering two elements of our reasons-responsive systems: our moral sentiments, and the kind of handling reasons seemed to call out for. It affects the former system by producing a normative outlook according to which normative properties are loud or powerful, and the normative landscape seems cosmic in scope or mystic in character. It affects the latter part of our reasons-responsiveness by altering what the agent takes to be an appropriate response to reasons. Mania gives rise to a sense of pervasive urgency and so the reasons that are marked for salience by the agent's reasons-tracking system are treated as if they were requirements. Proper recognition of this change involves recognising the manic sense of things as a genuine perspective on normative matters, which gives rise to real normative testimony. Maintaining the agentic stance towards manic people, rather than an explanatory stance (Mitova 2020) or objective

stance (Strawson 2008) is necessary to recognise the agential capacities at work when a manic person offers their act-description for something they have done (in the case of responsibility) or offers their view of reasons in any other regard (in the general course of being an agent). The second change is properly recognised when the way the manic agent handles reasons is recognised as responding, however imperfectly, to the normative properties of things. Celebrating the aesthetic properties of a painting, or ranting about the evils of some ancient transgression, might be inappropriate because they are done in a way that is out of balance with the reasons as they stand, but this is an imperfect, rather than failed, response to reasons. Abandoning plans in order to work on an inspired poem is often ill advised, and treats some of the agent's practical reasons too flippantly, but it also recognises the aesthetic value of this poem. The manic agent has found, and is still appreciating, the good aesthetic and bad moral properties of the objects of their fixations. Their perspective cannot be taken as default, however the context description according to which matters are urgent must be recognised as a legitimate initial position when considering how well or otherwise a manic agent responded to their practical reasons, subject to negotiation. Holding responsible, in the way I have outlined, respects the agency at work when manic people respond to reasons as if they were requirements, acknowledging the context-description under which matters seemed urgent and so this reason seemed to be a requirement, but also presenting the (often, but not always, more reliable) description under which the instead-of reasons had greater weight.

Holding responsible in this way recognises manic agential capacities for what they are. It incorporates the manic sense of things as legitimate moral testimony by accepting it as an initial negotiating position, and by acceding to different act-descriptions from which the act seemed more appropriate. It also introduces alternative act-descriptions which highlight relevant instead-of reasons that might have been dismissed or wrongly taken as pre-empted. Recognising manic agency on its own terms incorporates these features and is necessary if the manic person's agency is to be recognised for what it is, rather than be misrecognised as a damaged version of someone else's agency. The practice of negotiating act-descriptions outlined here ensures that it is manic people themselves, on their own terms, who show up on the other sides of interpersonal relationships, and do so under fitting descriptions. This is an instance of the broader kind of appropriate recognition of agency that is necessary to render manic agents' reasons-responsive capacities autonomous and is an ideal model for that kind of recognition.

Responsibility practices are governed by a norm of maintaining sites of mutual flourishing. Some degree of recognition of the kind of agency we have is necessary for this, but only in balance with other requirements. At the beginning of this enquiry, I noted that a great many descriptions are fitting for each of us, some optional, and some not. In addition to agents, we are organisms, political animals, aesthetic subjects, and yet more things. 'Agent' is a non-optional description, but it seems likely (if we have a hope of flourishing at all) that 'social being' is also non-optional. To excel as an agent involves our agential capacities being at work and enjoying social recognition, but that this recognition is social is not an accident. Non-optional descriptions like 'agent' and 'social being', it seems, overlap at the boundaries. To describe us as any kind of agent, at any level of specificity, is to create expectations of a particular ideal worth striving for. This ideal involves a social dimension – that of recognition. To flourish in social contexts requires that we are recognised for what we are, and an inescapable description of what we are is 'agent', narrowed to the specific kind of agency available to us. Recognising mad agency in the way required to confer autonomy on it is not only to allow someone to excel qua agent, it is also to include them in necessary dimensions of social flourishing.

§6.6 Conclusion

This chapter has argued for a view of responsibility as answerability which can include manic people in moral practices that maintain and repair interpersonal relationships as sites of mutual flourishing. Whilst the way we recognise reasons and the way we handle reasons while manic can be distinguished in some regards, appraisals of either are continuous, as appropriate handling is somewhat confined by the kinds of handling a reason seems to call out for. Changes in the kind of agency we enjoy, at any given time, will alter the kinds of response we have to offer and see as appropriate. In order to maintain a relationship, both parties must decide on the description under which they will deploy some action for the purposes of that relationship. This negotiation is conducted under a norm of facilitating flourishing. Consequently, the appropriateness of one agent's responses to reasons is a consideration, but so is the kind of agency they had with which to respond. If the latter is not included, that agent cannot be fully included in the relationship and so it is not a site of their flourishing as we would hope it to be. Crucially,

the conditions under which this relationship can be repaired if it has been damaged are also indispensable considerations.

Mad people cannot be included in this negotiation if their act-descriptions are rejected out of hand, or if they themselves are unwilling to entertain the perspective that their neurotypical interlocutors hold. To facilitate virtuous negotiations, an attitude of reconciliation is needed, by which we are invested in our friends' being at home in our relationship. This attitude is a sometimes effortful achievement, which we might secure by specifically activating basic empathy, and overcoming the assumption that our own normative outlooks are adequate or exhaustive. This is not always easy, and so the kinds of 'red flag' or 'timeout' signals that emerge or are stipulated in relationships become useful. This might also require both self-narration and self-management practices which continually direct one's attention to the non-manic perspectives that one has had, will have, or will have to interact with, as we discussed in Chapters 4 and 5.

Responsibility is typically directed at small, easily individuated actions (however it needn't exclusively be), and so allows for local instances of recognition of manic agency on its own terms in local contexts. The kind of negotiation outlined here is necessary for responsibility as answerability to function, but is also necessary to include manic agency in the scope of recognised capacities onto which the social status of autonomy can be conferred. It is largely uncontroversial that holding someone responsible expresses respect for someone in virtue of their agency, but it does more than this. It also promotes that agency to a social status that is necessary for the agent to approach the ideal of autonomy. Just as human agents fail to flourish if they are misrecognised in the relationships that are to be sites of that flourishing, agency does not flourish if it is not recognised on its own terms in the agent's social context. Recognising manic agency, on its own terms, is therefore crucial to rendering manic agency autonomous. Responsibility does not assume autonomy, but rather confers it.

Conclusion: Bipolar Autonomy is one Variation Among Many

This enquiry has had two objectives. The first has been to argue that BoRD does not necessarily limit or impair autonomy, but rather means that a distinctive kind of autonomy is the ideal of agency worth striving towards for bipolar agents. Implicit in these conclusions is a pluralist understanding of autonomy as an ideal. This is a more substantive claim than that different people will achieve autonomy in different ways. That claim is consistent with a unitary ideal being multiply realisable. Rather, the pluralist understanding involves the recognition that the standard of ideal will vary depending on an individual's existential reality. Not every situation we might find ourselves in will change what the ideal of agency worth striving for is for us, just as not every situation will change what our best all-things-considered reasons to act are. Our existential realities, however, are the sums and ranges of situations we have no choice but to face in our lives. Demonstrating this contextually sensitive plurality in autonomy as an ideal has been the second objective.

§7.1 Bipolar Autonomy is an Existent Variety of Autonomy.

There are many places where divots are taken out of our autonomy. By extension, there are many ways to be autonomous, and many shapes autonomy might take. The topographical project that Killmister embarked on (2017) invites us to think about autonomy in all its dimensions as something that can grow or diminish in many ways, that can be shaped and reshaped, and which might fit more or less well in the space available to it. Much of this investigation has served to offer proof by example that some kind of autonomy is available to bipolar agents, which accommodates their existential reality, and is consequently distinct in its shape and texture. The first part of this enquiry (Chapters 1-3) has been concerned with understanding the ways in which depression, mania, and repeated affective shifts alter the agentic capacities that are the foundation of this autonomy. Some of these capacities are made less reliable or less likely to be exercised, as when futility thinking blinds us to much of the value in life and leaves us unlikely to act on practical reasons that call on us. However, these shifts leave intact, and sometimes sharpen, other agentic capacities, such as those we use to really *appreciate* the normative properties of things. This is perhaps a less familiar response to a practical reason, and one that leaves few traces in the material world, but is nonetheless a response to reasons where they are apparent.

Relying on a different balance of familiar and unfamiliar agentic capacities, the bipolar agent must thread exercises together in a different way from neurotypical agents in order to produce a reliable diachronic agency. Consequently, their exercises of diachronic agency will be somewhat distinctive. Despite these differences, bipolar agency can still be the foundation of a robust personal autonomy. The distinctive features of that autonomy, its shape and texture, are inherited from the differences in the agency that underlies it.

§7.1.1 The Range of Responses to Reasons

Depression and mania can both reduce the rate at which an agent responds appropriately to practical reasons. Depression blinds us to much of the good, beauty, and meaning in the world, while stripping us of the motivational *oomph* to pursue any that is left. Mania presents some reasons as so urgent that considering any relevant instead-of reasons seems like an unjustifiable indulgence. On any reasons-responsive conception of agency, to respond to reasons less fully or reliably than they might otherwise have is an impairment. Depression and mania are impairments with respect to agency because they limit the extent to which an agent appropriately responds to reasons. Yet, they are not the only changes that BoRD produces in our agency, neither do they prevent that agency from grounding robust personal autonomy. Depression leaves intact minimal capacities to appreciate some normative values, particularly those which are grim and unpleasant, and mania sharpens this capacity to an exceptional degree. Appreciation of value is an agentic capacity. One way to respond to any value, the apprehension of some fact with an awareness of its normative properties, is to name it. Describing the normative properties of some fact, situation, or world, is a way of responding to it. This is a more minimal response, and one that is insufficient in the face of a great many reasons, but it is nonetheless a way of responding to reasons which exercises basic agential capacities. The futility thinker still resolves their experience into a normative outlook. In the testimonies we considered, it was clear that at least some still give moral testimony. This testimony demonstrated the self-answerability that Westlund takes to be crucial for autonomy, but is also a response to reasons (real or putative) in its own right. Despite these affective shifts, the capacity to describe the normative world is retained. Recognition of the depressed person's perspective is, therefore, recognition of some agential capacities at work and so is a necessary step to letting those capacities, such as they are, ground whatever autonomy is available. Pathologizing the manic person's frenetic ramblings, in contrast, dismisses the appreciative agential capacities at work and silences moral testimony.

Agency is impaired because our agentic capacities become less reliable against the standard of appropriate response, but what matters if these capacities are to ground our autonomy is that those standards continue to apply. BoRD impairs our agency in some regards, but the capacities that are impaired, that are left intact, or that are sharpened, all continue to count as agentic capacities, criticisable against the standard of appropriate response to reasons. The manic agent will often be mistaken about what response is appropriate, but we must note the difference between an *appropriate* response, which is justified or demanded by some particular reason, and a response that is *adequate*, that is, which counts as a response for the purposes of distinguishing reasons-responsive from non-responsive systems. Mania will often produce inappropriate, but nonetheless adequate, responses to practical reasons. The manic person, with their frenetic urgency, recklessness, and obsessiveness, will often fail *as* an agent, but not fail *to be* an agent. Thus, it is on this less than perfect agency that their autonomy can be founded.

Appreciative capacities respond to reasons and value, perhaps imperfectly or inappropriately, but adequately for the purposes of distinguishing agency from non-agency. The depressed agent might have to fall back on this as the only responsive system they have left, but this is not so for the manic agent. The affective shift towards mania changes how value is experienced, and so changes how agents encounter practical reasons. Like depression, mania changes the eyes with which we see the normative properties of things, and so value that is stark, grand, cosmic, or mystical is what is highlighted. But mania's influence is more pervasive, also changing what kind of response it is that reasons seem to call out for. Pervasive urgency is a restructuring of agency that changes the way in which agents handle and respond to reasons (actual or putative). This change might leave us frequently, perhaps even necessarily, less than fully successful in being a reasons-responsive agent, but in such cases we fail *as* agents, continuing to invite scrutiny against the reasons-responsiveness standard.¹⁰⁴ The manic agent takes reasons as demanding action and as pre-empting other considerations. To act, without further thought, is then to respond, in however imperfect a way, to one's apparent reasons.

¹⁰⁴ What we are discussing here is agency that is less than wholly effective, but which is still agency, and in many cases still at least somewhat effective. For a discussion of necessarily less-than-effective agency, see Radoilska, (2013, 2022).

§7.1.2 - Diachronic Agency

The first half of the investigation considered agential capacities in minute detail, but if these capacities permit us only to be a time-slice agent then it seems only an extremely local sort of autonomy could be available to us, if any at all. BoRD presents obstacles to the Bratmanesque planning agency that neurotypical agents might more readily pursue (Bratman 1999). For this reason, we considered the ways in which bipolar agents thread portions of their agential lives together in Chapters 3-5. The agential capacities used are sometimes different, or are deployed in different ratios from neurotypical agents, but as long as the bipolar agent can combine the discreet exercises of these capacities in a way that amounts to suitable diachronic agency, they have suitable foundations for their distinctive kind of personal autonomy.

There is more than one way to weave threads of agency together into a more cohesive diachronic whole, upon which less narrowly local aspects of autonomy are founded. Affective shifts can disrupt existing plans, but they can also leave an agent with a project half completed, that they did not realise had begun to unfold. Smaller projects can be woven together into larger projects, even when they had not initially been connected, particularly when the manic sense of things draws connections that would otherwise not be obvious to us. The exercise of agency during a manic episode at time t_1 counts as its own exercise, but can also be counted as the beginning of a longer diachronic exercise of agency, perhaps culminating long after that episode remits when the agent finishes some artwork at time t_2 . Deploying past activity as the beginnings of a project one finds oneself in the midst of is an exercise of wide-scope self-determination, but the variety in that mass of agency is provided by exercises of narrow-scope self-determination. Conducting one's agential project in this way requires a flexible narrative competency, but also requires that agents trust that their past exercises of agency can contribute to, or be part of, what they are doing now. Self-trust grants an agent the confidence to deploy past deeds as components of the longer projects they find themselves, unexpectedly, engaged in.

The bipolar agent, then, can be an opportunist, taking advantage of the rich and varied mass of activity behind them, and actualising projects out of emerging potentialities.¹⁰⁵ This opportunist mode relies on more narrow-scope, and less wide-scope, self-determination than more typical kinds of planning agency do, and founds autonomy with a distinctive texture (at least in its self-determination dimension). This difference in texture is something we can only capture when we are attentive to the difference in the scope at which self-determination is predicated, and capturing this difference in texture is necessary if we are to recognise some distinctive exercises of autonomy.

The diachronic agency that is available to bipolar agents lays different sorts of foundations from those laid by more neurotypical sorts of agency, and these differences transmit up to the shape and texture of the autonomy that is established. Exercises of self-determination (or of autonomy in any other dimension) inherit some of the features that mark out the exercises of diachronic agency they are established upon. Bipolar agents have at their disposal modes of agency which appreciate value or which opportunistically establish long term projects, more readily than do neurotypical agents, yet perhaps have less readily available the reliable planning agency that Bratman highlights. Insofar as this is the kind of agency available, the ideal of autonomy worth pursuing is not going to be the same as the ideal for neurotypical agents. The distinctiveness of bipolar autonomy is rooted in the kinds of agency that are available, given the instability of affective shifts, the need to scaffold agency across time, and the particular insights that extreme emotional experiences offer. This existential reality must be confronted in a different way, placing the kind of autonomy that is worth striving for at some distance from the ideal of autonomy that neurotypical agents ought to strive for.

¹⁰⁵ Recall Steward's mereological universalist account of agency as process, (2012). On such a view, our larger projects can be mereological sums of smaller projects or of unindividuated activity. Successive affective shifts provide a more varied mass of agency behind one, that can be retroactively deployed as the beginnings of the project one later commits to.

§7.2 Recognising Bipolar Autonomy

So far we've proven that the basics are in place. There are agential capacities that might be the foundations of autonomy, and they operate in ways that are somewhat distinctive. Thus, their ideal expression is somewhat removed from the ideal for agents confronting more conventional existential realities. What this means is that *if* bipolar autonomy exists as a non-impaired kind, it is dissimilar from neurotypical kinds in the ways described. It might be objected that the antecedent has not been demonstrated. Should we call this local ideal of agency 'autonomy'? In what key are we referring to it as autonomy?

The range of things we mean when we talk about autonomy is wide, and this has led some to question whether there is one concept of autonomy that we aim at while trading in different conceptions, or whether we have one word for a range of concepts, (Arpaly 2002, 117-148). Such disagreements have been interpreted to indicate that autonomy is a de facto contested concept (Radoilska 2012b, 252), and could be argued to be an *essentially* contested concept, (Radoilska 2021b, 149). If autonomy is essentially contested, then the satisfactory argument here is not one that persuades us that what we have outlined here is a kind of autonomy, but that it should be. This mirrors the adequate argument as to who are the *moral* champions out of our preferred teams, which does not hold that either has scored most goals or been the most sportsmanlike, but rather holds that whichever properties our teams have are the ones that count for determining the moral champions, (Gallie 1956). The argument to offer here is that which Sinclair suggests (1992) and which Rashed delivers in full (2019). Sinclair entreats us to: "[g]rant me the dignity of meeting me on my own terms – recognize that we are equally alien to each other, that my ways of being are not merely damaged versions of yours" (1992, 302). If autonomy is essentially contested, then we are called on to justify setting the boundaries in such a way that casts some people as defective agents, having only impaired autonomy. Contesting the boundaries is how we trade in essentially contested concepts, however extrinsic reasons can make some ways of contesting them more or less desirable than others. If we set and then contest the boundaries of autonomy in a way that is or might be exclusionary to bipolar agents, or other neuro-atypical people, then we expose this population to the dangers and harms of misrecognition that Rashed discusses (2019, 105-7). Mad agents who are subject to misrecognition suffer social exclusion and are frustrated in the formation of

their identity. These are important harms that significantly arrest people's flourishing and insofar as the consequences are produced by a community electing to trade in exclusionary conceptions of autonomy for the sake of using an essentially contested concept, is a matter of justice. That it is useful, in some sense, to trade in an essentially contested concept of autonomy in this way is not adequate to answer the charge levied by mad activists, because mad people are excluded from enjoying the utility generated when they are deployed as edge cases or negative examples.

Sinclair's entreaty, and the charge levied by the neurodiversity movement which it so concisely expresses, has normative force in contestations of autonomy, if autonomy is an essentially contested concept. If it is not essentially contested, however, but happens to be widely contested, much the same considerations have normative force partly grounded in epistemic, rather than directly political-moral values. If autonomy is contested, but non-essentially so, then some conception or conceptions of autonomy are what we are aiming at, and our contestations are valuable insofar as they get us closer to those conceptions. If some of the views that would make this contestation productive and bring us closer to the conceptions we are aiming at are shut out of our enquiry, then we have a problem. This is a problem for our understanding of autonomy, insofar as we are missing key perspectives that would inform a better view of autonomy. Chapman and Carel (2021) have presented a structurally similar argument pertaining to autism and flourishing. They argue that autistic people are routinely victims of testimonial injustice (see Fricker 2007, 9-59) when their reports that they are happy, thriving, or flourishing are not taken to be credible, or are taken as evidence that they have not very pronounced autism. They are victims of hermeneutic marginalisation (see Fricker 2007, 152) insofar as autistic perspectives on flourishing are shut out of the community's investigations on this matter. This is a moral wrong done to autistic people as a class, but it is also epistemically harmful to the community engaged in this enquiry. It is precisely because autistic people enjoy privileged epistemic access to knowledge about what autistic flourishing looks like that their testimony receives *deficient* credibility. Otherwise, it might be unjust to exclude autistic people from discussions about what flourishing is, but it would not be a testimonial injustice in particular. If autistic people were simply mistaken about how happy they are, the community would not wrong them specifically by ascribing low levels of credibility to this testimony. The same dynamic is at work with bipolar agents, who contest the boundaries of the autonomy concept. Bipolar agents are likely to recognise their own autonomy, familiar as they are with its particular lumps and contours, where neurotypical agents are not. The tendency to dismiss mad

perspectives on autonomy is rooted in the same prejudice as the epistemic injustice towards autistic people. The social identity of ‘mad person’ is both practically and epistemically unagenting. Mad people are taken not to be proper sources of practical and epistemic reasons in the way that neurotypical agents, especially those who enjoy other social privileges, are.¹⁰⁶ It is unjust to mad people to do this, but it is also unproductive for the purposes of enquiry. Bipolar agents can have privileged knowledge of the workings and details of their own kinds of agency, and of the kinds of autonomy that can be established upon it. If bipolar perspectives on agency and autonomy are not included in the non-essential contestation of the autonomy concept, key insights as to the workings of less typical kinds of agency are lost.

It could be objected that, whilst it would be unjust and epistemically counterproductive to exclude bipolar and other mad perspectives on autonomy from this productive contestation, they can be dismissed after consideration if they are found to describe something that is not autonomy. Perhaps what we have here called ‘bipolar autonomy’ is in fact some other phenomenon, perhaps adjacent to autonomy, but not an example of it. It is not possible to provide a satisfactory argument to the effect that what we have described is really autonomy without picking out at least a conception, if not a whole theory, of autonomy. I have kept commitments on both points as minimal as possible in order to allow that the outcome of this enquiry can be applied to a range of theories which differ in their conceptions of autonomy. The more specifically I respond to this objection, the less widely my conclusions apply across the field. Instead, I return to the role autonomy plays as an ideal of agency. Whichever capacities count as autonomy capacities on a given conception are capacities we are interested in because they are capacities-cum-value. Autonomy is an ideal of agency in the sense that it is the kind of agency we (should) hope to enjoy ourselves, and (should) hope that other agents enjoy. We have good reasons to support other agents in the development and protection of these agential capacities. Such reasons are central justifying concerns in some theories of political philosophy (Raz 1986, Colburn 2010, cf. Conly 2012), however we still have good reasons to prefer social arrangements that support more

¹⁰⁶ This terminology is taken from the central hypothesis of the *Epistemic Injustice, Reasons, and Agency* project (PIs Lubomira Radoislka and Veli Mitova). Underlying more familiar sorts of epistemic and practical injustice, it is suggested, is the fact that “unprivileged knowers’ reasons for action and belief are treated as being ‘of the wrong kind’ by the default. In other words, their reasons are conceived as explanatory at best, but not as potential justifications.” (EIRA 2018). I am indebted to the research network attached to this project for many fruitful discussions throughout the course of this investigation.

robust autonomy rather than less, all things being equal, even if we do not give autonomy this privileged place in political morality. When the boundaries of autonomy are contested, we are disagreeing as to what ideal of agency is worth striving for. If the existential reality that bipolar agents face makes it difficult or impossible to achieve the ideal of neurotypical agency, it need not follow that autonomy is closed to them. What we have demonstrated in this investigation is that another kind of agency can be found, and counts as the ideal local to the bipolar agent's situation. The bipolar agent has some agency, that might be expressed in a wide range of ways. The fact that they will likely not live as if they did not have BoRD does not mean that their agency is without value and that striving towards an ideal is no longer worthwhile. What is worth striving for is partly dependent on what avenues are available to us, what goes with the grain for us, and what we want. The bipolar agent, insofar as their agency reflects their neuro-atypical status, has reasons to strive for a kind of agency somewhat unlike that which their neurotypical peers should strive for. Yet there is still something worth striving for, and whatever occupies that space is the ideal of autonomy local to that agent. What counts as autonomy, therefore, is different for different agents. Recognising some behaviours as exercises of agency, as reaching out for that ideal, or as successfully exhibiting excellence in agency, is to recognise that behaviour as an exercise of autonomy. Consequently, what we have examined under the name 'bipolar autonomy' in this investigation is what we mean by autonomy as a (local) ideal. Deploying it sets a standard of aspiration and makes the exercise of some agential capacities count as exercises of autonomy-capacities.

§7.3 Implications

§7.3.1 Autonomy is Pluralist and Contextual

Many ways of arranging and exercising agential capacities might be sufficient to achieve some ideal of autonomy. In this sense, it has always been understood that autonomy is multiply realisable. What this enquiry has revealed is that the ideal itself is highly contextual, admitting of great plurality. There are not only different configurations and exercises of agential capacities that satisfy the ideal, but different ideals worth striving towards, partly depending on the configuration of capacities at an agent's disposal. The existential reality that an agent confronts has meaningful consequences for what condition of agency is worth their aiming for, and what condition their friends have reasons to support them in establishing. To speak of the shape and texture of someone's autonomy is a useful way of describing

these differences. If we understand agency as modular, as outlined in §2.4 and §3.6, we should expect most people's autonomy to be fairly lumpy. Different agential capacities, arranged in different orders, filling different roles, allow us to respond to the wide and contrary range of situations we will confront. For agents who confront sets of problems that most other agents do not, and which require particular modes of agency to respond to, they will likely adopt similar modes, which produce similar lumps and bumps in their autonomy. This is the difference that having an existential reality in common has for agents. Mad agents must confront problems that call for agentive solutions that their neurotypical peers do not require solutions to. Neurotypical agents might readily recognise autonomy when it is lumpy in a way familiar to them, but lumpiness admits of degrees. A potato is lumpy, but root ginger is lumpier. Mad agents who confront similar existential realities in common, will probably have lumps in their autonomy in common, familiar to each other but strange to those used to neurotypical kinds of autonomy.

The ideal condition of autonomy is indexed to one's starting position: what agential capacities are available or well tuned, what experiences will the agent likely confront, what options are open to them and what options are likely to remain closed. When standards are insensitive to our context, they can become perverse. Kennett and Wolfendale (2019) note the perversity of being expected to exercise high degrees of self-control when it will achieve little or nothing for one. The systemic racism and poverty that they consider are existential realities that change what agents should strive for. The ideal for someone who enjoys 'moral security' is a perverse standard for someone who does not. Agents who have unreliable capacities to conduct their agency in some particular way face a similarly perverse standard if the ideal of autonomy they are to strive for is insensitive to their capacities. One's existential reality is the sum of contexts, situations, and experiences that one will face which weigh on what ideal is worth striving for. Autonomy as an ideal condition, then, must be context-sensitive on pain of perversity. Avoiding that pain, there are a plurality of ideals of autonomy, more or less worth striving for depending on where one's starting position is.

§7.3.2 Autonomy Theory Must be Pluralist and Contextual

What does this conclusion mean for theorising about autonomy? What features should we expect new and emerging theories of autonomy to have if they are to be adequate to capture the plurality and contextual sensitivity of autonomy as an ideal?

One implication is that theories of autonomy should specify what starting point they take their candidates for autonomy to stand in. Theories should specify which agentive capacities in particular they have in mind, and in what configurations they envisage their being used. Previously it has been possible to be non-specific on this point, allowing that the multiple realisability of the ideal would absorb these differences. However, as we now recognise the degree to which the ideal is moved and altered by the capacities available to agents, knowing which capacities are in question becomes important for defining the boundaries of this ideal. To neglect this specification risks positing some condition as *the* ideal, rather than one among several, which obscures the variations in autonomy that it now behoves us to find. It will also no longer be useful to leave the agentive capacities in question as a free variable, to be filled in by the best philosophy of action, when deploying the concept of autonomy as an ideal, for instance, in political theory. Rather, we must know which subjects, identified by their kinds of agency, are the ones whose autonomy can be supported by these social arrangements or political institutions. Indeed, this greatly complicates the task of the autonomy-minded liberals or liberal-perfectionists, who take the objective of legitimate politics to be ensuring that more people can enjoy more autonomy. Given the plurality and contextual sensitivity at work, conflicts between kinds of autonomy are likely.¹⁰⁷ These conflicts might be manageable, there without one ideal of autonomy, it is more likely that measures to maximise access to the ideal of autonomy in fact maximise access to some ideals, at the expense of others whose local ideals are restricted. The universalist ambitions of such liberalisms are then returned to questions of distribution between competing citizens.

Agential capacities, as well as any other feature of an existential reality, will be important features of the starting position to note, so that we can identify the locale of any ideal of autonomy offered. More

¹⁰⁷ Further, supporting impoverished people in reaching the ideal of autonomy local to their condition cannot replace the political project of eliminating poverty. However, preparing them for autonomy after poverty is a perverse objective in situations where that poverty is not realistically going to be ameliorated.

useful than this, however, would be indications as to how this ideal might change if we take one step in any direction from that starting position. How does the ideal worth striving for change if someone's agentive capacities are altered in some way, if their existential reality becomes more harsh. What ideal of agency is worth striving towards, for instance, if we find ourselves in conditions of actual urgency? The ideal of (local) autonomy in such circumstances will not be the same as in more relaxed times. A comprehensive map of how an ideal of autonomy relates to its neighbours is of course beyond what could be expected of any one theorist or work, however indicators will be a useful feature of theories developed in recognition of this contextual sensitivity.

These are features we would hope newly emerging theories of autonomy might have, in respect of the plurality and contextuality of the ideal. However, a broader project is called for. This enquiry has offered an initial sketch of what is distinctive about bipolar autonomy, however a more robust account is required. Similarly robust accounts are justified for the bespoke sorts of autonomy that confront a range of similar and dissimilar existential realities. Schizophrenia and ADHD have some overlap with BoRD, but also present their own challenges that might call for distinct kinds of autonomy. Kennet & Wolfendale noted that self-control could be a perverse ideal in the face of poverty, but leave open the space to develop a positive account of what ideal(s) of agency are worth striving for from a position of moral insecurity, (not surrendering the project of eliminating that insecurity, but recognising how much progress in that project is likely to be made).

§7.3.3 Existing Theory Must be Critically Reconsidered

The kinds of autonomy that are treated as marginal cases or negative examples in existing theories are not properly understood if they are in fact different, but not deficient, cases of autonomy. What is called for is a widespread exercise in critical reconsideration. Exclusion of neuro-atypical agents is widespread, but not foundational to most existing theories, (cf Edward 1981). Critical reconsideration involves looking at each element of a theory individually and judging whether it does valuable work expanding our understanding of autonomy, or whether its work is to exclude atypical types of autonomy from understanding. It would be supremely wasteful to write off all hitherto developed theory of autonomy, and expend a great deal of effort rediscovering the obvious. Consequently, the portions of theories that are justified or motivated by putative negative examples, which we now recognise as

exercises of bipolar autonomy, can be removed and replaced, preserving the bulk of these theories. The conceptual tools developed during this enquiry will prove useful.

Appreciative agency responds to reasons in a way that does not produce changes in the material world which other agents might be able to rely on, but does respond in the sense that anger responds to injustice or gratitude responds to unexpected kindness. The depressed person who ruminates on the evils of the world or their own shortcomings is enacting some of the harmful features of depression, but they are doing so by exercising their appreciative agency. The appreciative mode of agency provides us the conceptual tools required to take an agentic stance towards these presentations where they would more usually be pathologized in depression and mania. This agentic stance is necessary in order to confer the social recognition onto agency (qua capacity, or qua any particular exercise of that capacity) that is necessary for it to become autonomous. Trading in terms of the appreciative mode of agency prompts us to recognise and include these distinctive depressive and manic expressions of agency. Such exercises count, then, as exercises of variant autonomy, rather than as negative examples. When reconsidered in this way, existing theories of autonomy are in a better position to accommodate bipolar autonomy.

This conceptual tool is likely to have broader application. The neo-Aristotelian tradition is rich with instances where the way in which we appreciate something invites praise or criticism, however appreciative capacities are more usually deployed as background features of agents, rather than as an exercise of agency in response to value. Thinking about appreciation in the latter terms might shed light on the relationship between value and agency, offering us a richer vocabulary with which to describe the interaction. There are metanormative implications to this thinking, which cast value as something that appeals to us rather than demands of us or sanctions us with the charge of irrationality. Exploring this relationship in more detail will also inform understanding of the role appreciation plays in ethical life. Srinivasan has suggested that it is an injustice to impose an invidious choice on someone between addressing an existing injustice and appreciating the wrongness of that injustice, (2018). Yet there are plausibly cases where we should not appreciate wrongs done to us fully, in order to facilitate forgiveness. The role our appreciative agency plays in these cases merits investigation.

We understood self-determination as being predicated at different scopes in different contexts. The scope at which self-determination is predicated changes the role self-determination will play in our life, and alters the textural qualities of our autonomy. If our understanding of self-determination is restricted to wide scope exercises, then we are likely to make this mistake, and will not recognise exercises of autonomy when they occur, nor will we be able to support bipolar (and other) agents in crafting autonomy for themselves that is adequate to confront their particular existential realities. To ask ourselves the question, ‘is this agent attempting to exercise autonomy over a different scope?’ is the useful question we can ask when considering self-determination, or any other dimension of autonomy. Asking this question directs our attention to the agent’s existential reality, and what the ideal of agency might look like for them given their situation. Applying this conceptual tool to existing and possible future theories of autonomy allows us to avoid the universalisation of one standard ideal of agency against which many of us will fail, and succeeding against which would be of only limited value. Understanding the way different agents zoom in and out when considering their agency, how it is divvied up into parts, and what role different parts are to play in striving for an ideal, better informs us about the details of their situation and the contours of the autonomy they are attempting to establish. As noted in §5.7, I suspect understanding the self as being predicated at scope might prove fruitful, particularly in questions of self-definition, self-narration, or self-creation. It is not obvious to me that the self will *always* be the largest contiguous who-shaped thing in one’s vicinity, (contra Schroeder 2022). It seems we often do, rightly, bring less than our entire selves to the table, when the scope at which our life is relevant narrows.

The final tool is the practice of co-stipulating act-descriptions, discussed in Chapter 6. Whilst theorists have few opportunities to negotiate act-descriptions with the characters in their thought experiments, a conceptual tool is available which offers similar utility. Considering how a mad agent might describe a putative failure of autonomy, in any case where it is deployed as a negative example, provides a check on the chauvinistic tendency. Theorists of course have the power to stipulate the intended project in such cases, but doing so can make the hypothetical case less like real cases of BoRD, or other mental disorder. Thus, integrating our own theoretical choices of act-description builds self-answerability (Westlund 2009, see §1.4.2 above) into the practice of theorycraft. Autonomy is a phenomenon that is enacted and lived by mad and neurotypical agents alike (and in some ways, unlike). We are answerable, in our practices of theorycraft, to the people whose lives we are seeking to understand. As

a conceptual tool, this co-stipulation involves mad perspectives on autonomy in the process of developing greater understanding of autonomy. Attention to act-descriptions has an honourable lineage in philosophy, but what the co-stipulation practice offers is a practical tool which can introduce some of the moral clarity, that philosophical theory aims at, into interpersonal interactions outwith the seminar room. As a social practice, it builds answerable intellectual virtues into daily interactions in a way that might roll out the progress made in ethical theory in practical ethical life.

These conceptual tools will be useful in some, however likely not all, instances of revising existing theories of autonomy. Theories of autonomy that are developed in the future can also take advantage of these tools to ensure that mad agency is recognised and neuro-atypical sorts of autonomy are not mistaken for non-autonomy. What this means, however, is that theories have significantly less space to deploy madness as a negative example. This is a restriction in how some theory of autonomy is argued for, but hedges out only theories for which neurotypicality would be partly constitutive of autonomy. The presumption of non-autonomy that attaches to madness more usually turns on the assumption that madness is very likely to injure some basic capacities needed for autonomy, but it is a stochastic, rather than comprehensive threat. Most theories would grant that sometimes, perhaps in rare cases, mad people are fortunate enough that their madness does not impair their capacities to an autonomy-defeating degree. These theories are restricted only in how they argue for their conclusion, not in their ability to reach it. Our enquiry has expanded those few exceptional cases into the wide possibility space of kinds of autonomy that can be maintained in the face of different existential realities, thus we have somewhat changed the explananda of theories of autonomy. Rather than seeking to explain why some general ideal of agency, (square and neurotypical as it might be) is what we should all strive for, we have demonstrated that the objective of theories of autonomy is to map the warren of agency, and understand what makes different ideals of autonomy worth striving for in different contexts.

Bibliography

Ahmed S (2004) *The Cultural Politics of Emotion*, London: Routledge

Alline, H (1806) *The Life and Journal of the Rev. Mr. Henry Alline*, Boston: Guilbert & Dean

Altshuler R (2014) The Value of Nonhuman Nature: A Constitutive View, *Ethical Theory and Moral Practice* 17(3):469-485.

Alvarez M (2010) *Kinds of Reasons*, Oxford: OUP

Anonymous, 'What is Bipolar Anger? An Insider's Perspective on Rage', *Speaking Bipolar*: [online] <https://speakingbipolar.com/what-bipolar-anger-rage/>

Last accessed, 7 Oct 2021

Anscombe G E M (2000) *Intention*, London: Harvard University Press

Aretaeus (1837) *of the Causes and Signs of Acute and Chronic Disease*, trans. T F Reynolds, London: William Pickering

Aristotle (2011) *The Eudemian Ethics*, Kenny A (trans) Oxford: Oxford World Classics

Arpaly N (2002) *Unprincipled Virtue: An Enquiry into Moral Agency*, New York: Oxford University Press

Bailey A (2018) On Anger, Silence and Epistemic Injustice, *Royal Institute of Philosophy Supplement* 84:93-115

Bailey O (2021) Empathy with vicious perspectives? A puzzle about the moral limits of empathetic imagination, *Synthese* (2021) 199:9621–9647

(2022) Empathy and the Value of Human Understanding, *Philosophy and Phenomenological Research*, 104:50-65

Bannon V B E (2017) Being a Friend to Nature: Environmental Virtues and Ethical Ideas, *Ethics, Policy and Environment* 20 (1):44-58

Beauchamp T L, Childress J F (2013) *Principles of Biomedical Ethics: seventh edition*, Oxford: OUP

Berman E P (2022) *Thinking Like an Economist*, Princeton: Princeton University Press.

Bierria A (2014) Missing in Action: violence, power, and discerning agency, *Hypatia*, 29 (1):129-145

(2020) Racial Conflation: Agency, Black Action, and Criminal Intent, *Journal of Social Philosophy*, 0(0):1–20

Binswanger L. (1933). *Über Ideenflucht*, [*Flight of Ideas*], Zurich: Fuesli

Bolton D, Banner N (2012) 'Does mental disorder involve a loss of personal autonomy?' in *Autonomy and Mental Disorder*, Radoilska L (ed), Oxford: Oxford University Press, pp77-99.

Bortolotti, L (2013) 'Rationality and Sanity: The Role of Rationality Judgements in Understanding Psychiatric Disorders'. In *The Oxford Handbook of Philosophy and Psychiatry*, Fulford K M W, Davies M, Gipps R, Graham G, Sadler J, Stanghellini G, Thornton T (eds), Oxford: Oxford University Press, pp 480–96.

(2015) The epistemic innocence of motivated delusions. *Consciousness & Cognition*, 33:490-499.

(2020) *The Epistemic Innocence of Irrational Beliefs*, Oxford: Oxford University Press

& Cox R, Broome M, Mameli M (2012) 'Rationality and self-knowledge in delusion and confabulation: implications for autonomy as self-governance' in *Autonomy and Mental Disorder*, Radoilska L (ed) Oxford: Oxford University Press, pp100-122

Brahim L O (2018) Reconsidering the 'self' in self- management of chronic illness: Lessons from relational autonomy, *Nursing Enquiry*, 2019;26:e12292

Brandenburg, D. (2017). *The Nurturing Stance: Making Sense of Responsibility without Blame*, *Pacific Philosophical Quarterly*, 99 (1)5–22.

& Strijbos D (2020). The Clinical Stance and the Nurturing Stance: Therapeutic Responses to Harmful Conduct by Service Users in Mental Healthcare, *Philosophy, Psychology, & Psychiatry*, 27 (4): 379–394.

Bratman M E (1999) *Intention, Plans, and Practical Reason*, Stanford: CSLI Publications

(2003) *Autonomy and Hierarchy*, *Social Philosophy and Policy*, 20(2), 156-176

Calhoun C. (2009). "Losing One's Self" in *Practical Identity and Narrative Agency*, Mackenzie C, Atkins K, (eds) London: Routledge, pp193-211.

Campbell J (1968) *The Hero with a Thousand Faces*, Princeton: Princeton University Press

Carter A J, (2020), Epistemic Perceptualism, Skill, and the Regress Problem, *Philosophical Studies*, 177:1229-1254

Chapman R, Carel H (2022) Neurodiversity, Epistemic Injustice, and the Good Human Life, *Journal of Social Philosophy*, 2022(00):1-18

Charland, LC (2005) 'Emotional Experience and the Indeterminacy of Valency' in Barrett L F, Niedenthal P M, Winkelman P, (eds) *Emotion and Consciousness*, London: the Guilford Press, 221-254

Cheney T, (2008) *Manic: a memoir*, HarperCollins ebooks

Christensen D (2009) Disagreement as Evidence: The Epistemology of Controversy, *Philosophy Compass*, 4(5): 756–767.

(2010) Higher-Order Evidence, *Philosophy and Phenomenological Research*, 81 (1): 185–215.

Coady D, (2017) "Epistemic Injustice as Distributive Injustice" in *The Routledge Handbook of Epistemic Injustice*, Oxford: Routledge, pp61-68

Colburn B (2010) *Autonomy and Liberalism*, London: Routledge

Conly S (2012) *Against Autonomy: Justifying Coercive Paternalism*, Cambridge: Cambridge University Press

Cowan R, (2016), Epistemic Perceptualism and Neo-Sentimentalist Objections, *Canadian Journal of Philosophy*, 46(1)59-81

Crichton P, Carel H, Kidd I J (2017) Epistemic Injustice in Psychiatry, *BJPsych Bulletin* 41:65-70

Damasio, A (1996) *Descartes' Error: Emotion, Reason, and the Human Brain*, Oxford: McMillan Publishers Ltd

Depp C A, Stricker J L, Zagorsky D, Goodale L C, Eyler L T, Patterson T L, Lebowitz B D, Jeste D V (2009) Disability and Self-Management Practices of People with Bipolar Disorder: a Web-based survey, *Community Mental Health Journal*, 45:179–187.

Dings R, Glas G (2020) Self-Management in Psychiatry as Reducing Self-Illness Ambiguity, *Philosophy, Psychiatry, & Psychology*, 27 (4): 333–347.

Dworkin G (1988) *The Theory and Practice of Autonomy*, Cambridge: Cambridge University Press

Edward R E (1981) Mental Health as Rational Autonomy, *The Journal of Medicine and Philosophy*, 6:309-322

EIRA (2018) *Epistemic Injustice, Reasons, and Agency*, (collaborative research project) homepage available online at: <https://epistemic-injustice-reasons-agency.weebly.com/>

Last accessed, 25 Nov, 2022, 1459 hrs.

Eliade M (1964) *Shamanism: Archaic Techniques of Ecstasy*, Trask W R (trans), Arkana

Feinberg J (1989) 'Autonomy' in *The Moral Limits of Criminal Law: Harm to Self* vol 3, Oxford: Oxford University Press, pp 27-51

Fernandez A V (2014) Reconsidering the Affective Dimension of Depression and Mania: towards a phenomenological dissolution of the paradox of mixed states, *Journal of Psychopathology*, 20:414-422

Fisher C (2009) *Wishful Drinking*, London: Simon & Schuster

Fischer J M (2012a) Responsibility and Autonomy: the problem of mission creep, *Philosophical Issues*, 22(1):165-184

(2012b) *Deep Control: Essays on Free Will and Value*, Oxford: OUP

& Ravizza M (1998) *Responsibility and Control: A Theory of Moral Responsibility* Cambridge: Cambridge University Press

Foot P (2001) *Natural Goodness*, New York: Oxford University Press

Frank A (2013) *The Wounded Storyteller: Body, Illness & Ethics*, London: University of Chicago Press

Frankfurt H (1998) *The Importance of What We Care About*, Cambridge: Cambridge University Press.

Frankl V E (1986) *The Doctor and the Soul: from psychotherapy to logotherapy*, Winston R, Winston C (trans), New York: Vintage Random House Publishing

Fricker M (2007) *Epistemic Injustice: Power and the Ethics of Knowing*, Oxford: Oxford University Press

Gallie W B (1956) Essentially Contested Concepts, *Proceedings of the Aristotelian Society*, 56:167-98

Gorman A (2020) Depression's Threat to Self-Governance, *Social Theory and Practice*, 46(2):227-297

Govier T (1993) Self-trust, Autonomy, and Self Esteem, *Hypaitea* 8(1):99-120

Harding S (2004) *The Feminist Standpoint Theory Reader: Intellectual and Political Controversies*, London: Routledge

Hartsock N (1987) "The Feminist Standpoint: Developing the Ground for a Specifically Feminist Historical Materialism," in *Feminism and Methodology: Social Science Issues*, Harding S (ed), Bloomington: Indiana University Press

Hieronymi P (2013) Reflection and Responsibility, *Philosophy and Public Affairs*, 42(1): 3–41

Hippocrates (2015) *Complete Works*, Hastings: Delphi Publishing Ltd

Holton R (1999) Intention and Weakness of Will, *Journal of Philosophy*, 92(5):241-262.

Houlders J W, Bortolotti L, Broome M R (2021) Threats to epistemic agency in young people with unusual experiences and beliefs, *Synthese* 199:7689–7704

Hursthouse R (2007) “Environmental Virtue Ethics” in *Working Virtue: Virtue Ethics and Contemporary Moral Problems*, Walker R L, Ivanhoe P J (eds), Oxford: Clarendon Press, pp155-172

Jaggar A (1989) Love and Knowledge: Emotion in Feminist Epistemology. *Inquiry*, 32(2), 151–76

James W (2002) *Varieties of Religious Experience: a study in human nature (centenary edition)*, London: Routledge

Jamison K R (1994) *Touched with Fire*, New York: Simon & Schuster

(1995) *An Unquiet Mind: A Memoir of Moods and Madness*, London: Picador

(2004) *Exuberance: The Passion for Life*, New York: Vintage Books

Jaworska A (1999) Respecting the Margins of Agency: Alzheimer’s Patients and the Capacity to Value, *Philosophy and Public Affairs*, 28 (2):105-138.

(2007) Caring and Full Moral Standing, *Ethics*, 117 (3):460-497

(2008) 'Caring, Minimal Autonomy, and the Limits of Liberalism' in *Naturalised Bioethics: Towards Responsible Knowing and Practice*, Lindemann H, Verkerk M, Walker M U (eds), Cambridge: Cambridge University Press, pp80–105.

Jones K (2003) Emotion, Weakness of Will, and the Normative Conception of Agency, *Royal Institute of Philosophy Supplement*, 52:181-200

Jordan J (2020) 'I was cut off from the world, bound in one place by illness' Interview: Susana Clarke, *The Guardian*, 12 September 2020,
<https://www.theguardian.com/books/2020/sep/12/susanna-clarke-i-was-cut-off-from-the-world-bound-in-one-place-by-illness>

Last accessed: 31 Oct 2022, 1218 hrs

Kahane G (2017) If Nothing Matters, *Noûs*, 51(2):327–353

Kennett J, Wolfendale J, (2019) 'Self-Control and Moral Security' in: *Oxford Studies in Agency and Responsibility Vol 6*, ed Shoemaker D, Oxford: Oxford University Press

Killmister S (2017) *Taking the Measure of Autonomy: A four-dimensional theory of self-governance*, Oxford: Routledge

Kirchin S (2013) 'Thick Concepts and Thick Descriptions' in *Thick Concepts*, Kirchin S (ed), Oxford: Oxford University Press, pp60-77

(2017) *Thick Evaluation*, Oxford: Oxford University Press

Korsgaard C (1996) *The Sources of Normativity*, Cambridge: Cambridge University Press

Laidlaw J (2014) *The Subject of Virtue: an anthropology of ethics and freedom*, Cambridge: Cambridge University Press

Lake C R (2008) Disorders of thought are severe mood disorders: the selective attention defect in mania challenges the Kraepelinian dichotomy – a review. *Schizophrenia Bulletin*, 34(1):109–117

Langton R (1993) Speech Acts and Unspeakable Acts, *Philosophy & Public Affairs*. 22(4):293–330.

Lodge P (2020) What Is It Like To Be Manic? *The Oxonian Review*, 3 Feb, 2020:

<https://oxonianreview.com/articles/what-is-it-like-to-be-manic>

Last accessed 7 Dec 2022

Longden E (2013) The Voices in my Head, *TED Conference 2013*, available online at:

https://www.ted.com/talks/eleanor_longden_the_voices_in_my_head?language=mn

Last Accessed: 22 Nov 2022

MacIntyre A (1985) *After Virtue*, London: Duckworth

MacKenzie C (2014) 'Three Dimensions of Autonomy: A Relational Analysis', in *Autonomy, Oppression, and Gender*, Veltman A, Piper M (eds), Oxford: Oxford University Press, pp15-41

Mariqueo-Russell A (manuscript) Desire and Motivation in desire theories of motivation,

Martin W, Gergel T, Owen G S (2019) Manic Temporality, *Philosophical Psychology*, 32:1:72-97

McGeer V, Pettit P (2015) 'The Hard Problem of Responsibility', in *Oxford Studies in Agency and Responsibility vol 1*, Shoemaker D (ed), Oxford: Oxford University Press

McGrath S (2019) *Moral Knowledge*, Oxford: Oxford University Press

McKay R, Langdon R, Coltheart M (2005) Sleights of Mind: Delusions, defences, and self-deception, *Cognitive Neuropsychiatry*, 10(4):305-326

McLeod C, Sherwin S (2000) 'Relational Autonomy, Self-Trust, and Health Care for Patients Who Are Oppressed', in *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self*, Mackenzie C and Stoljar N (eds), Oxford: Oxford University Press

McNamara J, DuBrul A, et al (2007) *Navigating the Space Between Brilliance and Madness: A Reader and Roadmap of Bipolar Worlds*, New York: The Icarus Project

Mitchell J (2021) *Emotions as Feeling Towards Value: a theory of emotional experience*, Oxford: Oxford University Press

Mitova V (2020) Explanatory Injustice and Epistemic Agency, *Ethical Theory and Moral Practice*, 23:707–722

Moore A, Hope T, Fulford K W M (1994) Mild Mania and Wellbeing, *Philosophy Psychiatry, & Psychology*, 1(3)165–177.

Moynihan J (2016) *Healing Voices*, dir Moynihan J,

Nguyen C T (2019) Games and the Art of Agency, *Philosophical Review*, 128(4):423-462

(2020) *Games: Agency as Art*, New York: Oxford University Press

Oshana M A L (2002) The Misguided Marriage of Autonomy and Responsibility, *Journal of Ethics*, 6:261-280

(2006) *Personal Autonomy in Society*, Aldershot: Ashgate

Parfit D (1984) *Reasons and Persons*, Oxford: Oxford University Press

Parker R (1983) *Miasma: Pollution and Purification in Early Greek Religion*, Oxford: Oxford University Press

Paul L A (2014) *Transformative Experience*, Oxford: Oxford University Press.

Petrolini V (2021) Too Much or Too Little? Disorders of Agency on a Spectrum, *European Journal of Analytic Philosophy*, 16(2):79-100

Pettigrove G (2007) Forgiveness and Interpretation, *Journal of Religious Ethics*, 35(3):429-452

Pickard H (2011) Responsibility without Blame: Empathy and the Effective Treatment of Personality Disorder, *Philosophy, Psychiatry, & Psychology*, 18(3):209-223

(2017) Responsibility without Blame for Addiction, *Neuroethics*, 10:169–180

Plato (2002) *Phaedrus*, Waterfield R (trans), Oxford: Oxford University Press

Potter N (2013) Narratives Selves, Relations of Trust, and Bipolar Disorder, *Philosophy, Psychiatry, & Psychology*, 20(1):57-65

Pugh J (2020) *Autonomy, Rationality, and Contemporary Bioethics*, Oxford: Oxford University Press

Radden J (1994) Second Thoughts: Revoking Decisions Over One's Own Future, *Philosophy and Phenomenological Research*, 54(4):787-801

(2013) The Self and its Moods in Depression and Mania, *Journal of Consciousness Studies*, 20(7-8):80-102

Radoilska L V (2012a) 'Introduction: personal autonomy, decisional capacity, and mental disorder' in 'Autonomy and Mental Disorder', ed L. Radoilska. Oxford: Oxford University Press, pp ix-xli.

(2012b) 'Autonomy and Ulysses Arrangements' in *Autonomy and Mental Disorder*, Radoilska L (ed), Oxford: Oxford University Press, pp252-280.

(2013) *Addiction and Weakness of Will*, Oxford: Oxford University Press

(2021a) Circumstance, Answerability, and Luck, *The Monist* 104 (2): 155-167

(2021b) 'Autonomy and Responsibility' Chapter 13 in *The Routledge Handbook of Autonomy*, ed Colburn B, London: Routledge, pp147-157

(2022) 'Pathologies of Agency' in *the Routledge Handbook of Philosophy of Agency*, Ferro L (ed), London: Routledge, pp169-177

Rashed M A (2019) *Madness and the Demand for Recognition: A Philosophical Inquiry into Identity and Mental Health Activism*, Oxford: Oxford University Press

Ratcliffe M (2010) Depression, Guilt and Emotional Depth, *Inquiry*, 53(6):602-626

(2013) What is it to Lose Hope?, *Phenomenology and the Cognitive Sciences* 12:597-614

(2020) 'Existential Feelings' in *the Routledge Handbook of Phenomenology of Emotions*, Oxford: Routledge, pp248-261

Raz J (1986) *The Morality of Freedom*, Oxford: Oxford University Press

Reid M (2012) 'I choose fiercely to live, but only for now.' *The Times*:
<https://www.thetimes.co.uk/article/i-choose-fiercely-to-live-but-only-for-now-stk2b8hxx8v>, last
accessed: 27 April 2021.

Sandis C (2015) *Action in Life and Art*, lecture to the Royal Institute of Philosophy, available at:
<https://www.youtube.com/watch?v=616F97QhDX0>

Last accessed: 15 Nov 2022, 1457 hrs.

Saas L, Pienkos E (2015) Beyond Words: linguistic experience in melancholy, mania, and schizophrenia,
Phenomenology and the Cognitive Sciences, 14:475–495

Scanlon T (2008) *Moral Dimensions: Permissibility, Meaning, Blame*, Cambridge: Harvard University
Press

Scullion S (1998) Dionysus and Katharsis in "Antigone", *Classical Antiquity* 17(1): 96–122

Schechtman M (2001) Empathic Access: The Missing Ingredient in Personal Identity, *Philosophical
Explorations*, 4(2):95-111

Schramme T (manuscript) Understanding people, empathy, and hermeneutical injustice.

Schroeder M (2022) Narrative and Personal Identity, *Aristotelian Society Supplementary Volume*
96(1):209-226

Schulman S (2016) *Conflict is Not Abuse: overstating harm, community responsibility, and the duty of
repair*. Vancouver: Arsenal Pulp Press

Sherdell L, Waugh C E, Gotlib I H (2012) Anticipatory Pleasure Predicts Motivation for Reward in Major Depression, *Journal of Abnormal Psychology*, 121(1):51–60

Shpall S (2014) Moral and Rational Commitment, *Philosophy and Phenomenological Research*, 88(1):146-172

Shoemaker D (2015) *Responsibility from the Margins*, Oxford: Oxford University Press

Sinclair J, (1992) “Bridging the Gaps: An Inside-Out View of Autism (Or, Do You Know What I Don't Know?)” in *High Functioning Individuals with Autism*, Schopler E, Mesibov G B (eds), New York: Springer Science+Business Media LLC, pp294-302

Solomon A (2001) *The Noonday Demon: an atlas of depression*, London: Vintage Random House

(2008) Depression, too, Is a Thing with Feathers, *Contemporary Psychoanalysis*, 44(4):0010-7530

Spaid A (2020). *Desire satisfaction theories and the problem of depression*. PhD. University of Nebraska.

Srinivasan A (2018) The Aptness of Anger, *Journal of Political Philosophy*, 26(2):123-144

Steward H (2012) Action as Process, *Philosophical Perspectives* 26(1):373-388.

Stocker M (1979) Desiring the Bad: An Essay in Moral Psychology, *The Journal of Philosophy*, 76(12):738-753

Strawson P (2008) *Freedom and Resentment and Other Essays*, Oxford: Routledge

Summers J S (2017) Post hoc ergo propter hoc: some benefits of rationalization, *Philosophical Explorations*, 20(1):21-36

Taleb N N (2012) *Antifragile: things that gain from disorder*, London: Penguin

Tappolet C (2016) *Emotions, Values, and Agency*, Oxford: Oxford University Press

Tekin Ş (2011) Self-concept through the Diagnostic looking glass: narratives and mental disorder, *Philosophical Psychology*, 24(3):357-380

Tenenbaum, S (2003) "Accidie, Evaluation, and Motivation" in *Weakness of Will and Practical Identity*, Stroud S, Tappolet C (eds) Oxford: Oxford University Press

- (2006) *Appearances of the Good: an Essay on the Nature of Practical Reason*, Cambridge: Cambridge University Press

Tredway M T, Zald D H (2011) Reconsidering Anhedonia in Depression: Lessons from Translational Neuroscience, *Neuroscience and Biobehavioural Reviews*, 35(3):537–555

Tully I (2017) Depression and the Problem of Absent Desires, *Journal of Ethics and Social Philosophy*, 11(2):1-16

van Willigerburg T, Delaere P, (2005) Protecting Autonomy as Authenticity Using Ulysses Contracts, *Journal of Medicine and Philosophy*, 30:395–409

Vellman J D (1992) The Guide of the Good, *Noûs*, 26(1):3-26

Walker M U (2007) *Moral Repair: Reconstructing Moral Relationships After Wrongdoing*, Cambridge: Cambridge University Press

Wallace R J (1994) *Responsibility and the Moral Sentiments*, Cambridge: Harvard University Press

Waterfield R (2002) 'Explanatory Notes' in *Phaedrus*, Plato, Waterfield R (trans), Oxford: Oxford University Press

Watson G (1975) Free Agency, *The Journal of Philosophy*, 72(8):205-220

Weiner T (2011) The (Un) managed Self: Paradoxical Forms of Agency in Self-Management of Bipolar Disorder, *Culture, Medicine, and Psychiatry*, 35:448–483.

Westlund A (2009) Rethinking Relational Autonomy, *Hypatia*, 24(4)26-49

Williams, (1981) *Moral Luck*, Cambridge: Cambridge University Press.

Wolf S (2015) *The Variety of Values*, Oxford: Oxford University Press

Whyte C (2002) 'Introduction' in *Dain do Eimhir: Poems to Eimhir*, by Somhairle MacGill-Eain, Whyte C (ed), Edinburgh: Polygon

WHO (2019) *International Categorisation of Diseases volume 11*, available online at:

<https://icd.who.int/browse11/l-m/en>

Last accessed 7 Oct 2021

Zahavi D, Fernandez A V, (2020) Basic empathy: Developing the concept of empathy from the ground up, *International Journal of Nursing Studies*, 110,

DOI: 10.1016/j.ijnurstu.2020.103695