ETHICS FOR THE DEPRESSED

A VALUE ETHICS OF ENGAGEMENT

by

DEVIN ROBINSON FITZPATRICK

A DISSERTATION

Presented to the Department of Philosophy
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2022

DISSERTATION APPROVAL PAGE

Student: Devin Robinson Fitzpatrick

Title: Ethics for the Depressed: A Value Ethics of Engagement

This dissertation has been accepted and approved in partial fulfilment of the requirements
for the Doctor of Philosophy degree in the Department of Philosophy by:

Mark Johnson	Chair
Colin Koopman	Core Member
Beata Stawarska	Core Member
Nicholas Allen	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2022.

DISSERTATION ABSTRACT

Devin Robinson Fitzpatrick

Doctor of Philosophy

Department of Philosophy

June 2022

Title: Ethics for the Depressed: A Value Ethics of Engagement

I argue that depressed persons suffer from "existential guilt," which amounts to a two-part compulsion: 1) the compulsive assertion or sense of a vague and all-encompassing or absolute threat that disrupts action and intention formation, and 2) the compulsive taking of such disruption to be a reason for inaction. I develop in response an "ethics for the depressed," an ethical theory directed to those suffering from existential guilt.

The first part of this dissertation, comprising Chapters 2 through 4, largely concerns the first aspect of existential guilt: it is a metaethics for the depressed, or "ethics as a reliable guide" as a response to "demoralization" and "hypermoralized deliberation." There I challenge what I call the Stocker-Smith account of depressive loss of motivation as being a loss of desires and argue instead that it involves the defeating presence of what the phenomenologist Matthew Ratcliffe calls "pre-intentional" mental states, a category that I redefine and expand to include second-order "quasi-beliefs" and habits of feeling, that interfere with intention formation and action despite the persistence of desire.

The second part of this dissertation, comprising Chapters 5 and 6, largely concerns the second aspect of existential guilt: it is a normative ethics for the depressed, or a "value ethics of engagement" premised on "contingent value ranking." After

demonstrating in the first part that depressed persons may retain their desires and values in depression, I premise a value ethics upon what I call the consistent desire for a "sense of stability" in response to experiences of precarity and isolation. From the phenomenology of value, I develop a concept of the heart as the set of "felt values" or intuitive value paradigms that are themselves pre-intentional states or dispositions.

I thus attempt to structure a complete ethical theory, integrating plural philosophical traditions and founded on the phenomenological category of pre-intentional mental states, in response to the presence of existential guilt and its component compulsions as experienced by an otherwise reasonable interlocutor. I put an orthodox style of philosophy in service of an unorthodox agent: one who is "aspiringly autonomous."

CIRRICULUM VITAE

NAME OF AUTHOR: Devin Robinson Fitzpatrick

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
New School for Social Research, New York, NY
University of California, Berkeley

DEGREES AWARDED:

Doctor of Philosophy, Philosophy, 2022, University of Oregon
Master of Arts, Philosophy, 2014, New School for Social Research
Bachelor of Arts, Philosophy, 2008, University of California, Berkeley

AREAS OF SPECIAL INTEREST:

Ethics and Value Theory
20th-Century Continental (especially Phenomenology)
20th-Century American (especially Pragmatism)

GRANTS, AWARDS, AND HONORS:

2019 Institute of Cognitive and Decision Sciences Award ($5,400)

Oregon Humanities Center's Graduate Research Support Fellowship ($1,000)

University of Oregon Graduate School's Special "Opps" Travel & Research
Award ($1,000)

Accepted to Copenhagen Summer School in Phenomenology and Philosophy of
Mind 2016 at the Center of Subjectivity Research, University of
Copenhagen

Accepted to Collegium Phaenomenologicum, 41st Annual Session, Summer 2016

University of Oregon Graduate Teaching Fellowship, 2014-2020

New School for Social Research Merit Tuition Scholarship, 2011-2014

PUBLICATIONS:

Articles and Book Chapters

> 2022. "Scheler's Phenomenology of Value as Value Pluralism," in *Max Scheler: His Thought and Influence*, edited by Susi Gottloeber, 221–38. Springer.

> 2014. "Realism and Receptivity: The Role of the Transcendent in Pragmatism," *Transactions of the Charles S. Peirce Society* 50 (1).

Book Reviews and Editorials

> 2018. Review of M. Baghramian, S. Marchetti (eds.), *Pragmatism and the European Traditions: Encounters with Analytic Philosophy and Phenomenology Before the Great Divide*. Phenomenological Reviews.

> 2018. "Editors' Introduction: Reflections on the First Issue," with Martina Ferrari, Sarah McLay, Shannon Hayes, Kaja Jenssen Rathe, and Amie Zimmer. *Puncta: Journal of Critical Phenomenology* 1 (1).

> 2018. Review of *Space, Time, and Other: A Study in the Method and Limits of Transcendental Phenomenology*, by Fred Kersten. Graduate Faculty Philosophy Journal 39 (1).

ACKNOWLEDGMENTS

I would not have been able to complete this work without the support of my wife, my family, my friends, my colleagues, and my committee. I am deeply grateful to you all. I also benefited from weekly talk therapy and anti-anxiety medication. I am grateful for this, too.

But I confess that most of all, I am grateful to say that I think I have changed. I used to believe that I was passionate but incurious. I had an urgent problem to solve, the question of how to live my life, and even my own interests were trivial in comparison. This understanding of myself was grandiose and, worse still, sincere. I tried to not inflict this fervor on others, feeling it would be unfair to judge them as I judged myself. I tried to resist allowing my contempt toward myself to become contempt toward others. I hope that I succeeded.

This dissertation mirrored my grand quest, and I despaired at its reflection. The closest I had ever been to solving my problem still felt infinitely distant. The best answers I had ever devised felt infinitely inadequate to the question. This has been my best opportunity yet to face my deepest challenges with my utmost abilities. Yet it is also an academic requirement for a doctorate, every excess page sure to try my committee's patience. This muddling of the transcendent and the mundane may be the essence of academia. Or it may just be what feels like the essence of academia to one who has lived in academia, because it is just life.

To write this, I had to change myself. I had to stop moralizing about my dissertation, working myself up into thinking I would save the world with a theory and cutting myself back down with shame and doubt. I had to experience the prospect of

failure as an acceptable risk. I had to be able to believe that it was worth doing my best at what could only be, at best, a steppingstone for something greater. Maybe again, too. Maybe again and again, forever.

Have I succeeded? Well, here it is.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I: INTRODUCTION

> "It is a mighty heritage, it is the human heritage,
> and it is all there is to trust. And I learned this through descending,
> as it were, into the eyes of my father and my mother."
> James Baldwin (2008, 60)

One symptom of depression is to find it difficult to begin. Another, or maybe a more nuanced interpretation of the same, is to find it difficult to see what one is doing as beginning. It's not strange to feel unworthy of a task, and it's not impossible to take on or complete a task of which one feels unworthy. But for a depressed or anxious person, not only is feeling oneself to be deeply unworthy a constant condition, that felt unworthiness seems to make every task impossible. It doesn't help that these tasks can be ambitious. After all, given how deep this unworthiness feels, only the highest possible ambitions promise redemption from it. And if those ambitions are impossibly high, then that suits, even justifies, the sense that the task, like every task, is impossible. If it is impossible, then every beginning is unworthy of it. To see oneself as beginning, then, already means to see what one is doing as unworthy, which is a single quick step, or slip, from being struck by one's own unworthiness. All this makes it difficult to begin. But, as you can see, even with an ambitious task, that difficulty is surmountable. The trick is in going on.

This work, the task of writing before me and reading before you, is ambitious. It attempts to justify an ethical theory that advocates beginning ethical deliberation with intuitions about what is valuable, even though such intuitions are unreliable. It critiques a dominant alternative, procedural ethics, and the ideal of objectivity to which such ethics tend to aspire. It is about what makes an ethical theory credible, and so it is about the all-too-often imaginary audience ethical theory tends to be for: those who are rationally

1

persuadable, skeptics or egoists, and not those who suffer from compulsive thinking that is resistant to argument, like the anxious or the depressed. It is an ethical theory for depressed persons, not the skeptic or egoist, and so it is about the problems distinct to excessive concern with ethical worthiness and the nature of all-encompassing guilt. It generalizes from the case of depressed persons to draw broader lessons about demoralization, the nature of ethical motivation, and how improperly structured ethical thinking – including theories – can cause pain to the depressed and demoralized. And, in being for depressed persons, it takes exhortation, not just explanation and justification, to be central to its success as a theory: it seeks not only to tell you how you are permitted to go on, but to induce you to feel able to go on. Even so, it seeks to avoid conflating theory with therapy, while reflecting on how best to incorporate therapeutic goals into ethical theory and the implications for evaluating ethical theories.

These matters seem too big to approach from the front, so I will approach them from the side. To introduce the terms with which I would present this ethical theory, the context from which it emerges, and the significance of both, I start with an example not of depression but of alienation: the case of a professional ethicist's complaint.

1. Personal and Impersonal Ethics

"There's a kind of thinking that we do when we are trying to prove something," Michael Sigrist writes, "and then a kind of thinking we do when we are trying to do something or become a certain kind of person—when we are trying to forgive someone, or be more understanding, or become more confident in ourselves." The quote is taken from Sigrist's 2019 American Philosophy Association blog post "Why Aren't Ethicists

2

More Ethical?," in which he reflects on Eric Schwitzgebel's provocative 2011 paper claiming that professional ethicists are no more ethical than average.[1] Sigrist glosses the usual objections to, or justifications for dismissing, Schwitzgebel's claims: that morality cannot be measured or, if it could be, we'd first have to know the proper object of measurement – intentions, consequences, character – about which deontologists and consequentialists and virtue ethicists all disagree. The "findings ring true to me," Sigrist instead emphasizes, and "[i]t's the truth of this claim—not how it is established—that bothers me." Sigrist's conclusion is that this truth lies in the different kinds of thinking involved in ethical theory and in practical efforts to live more ethically: where the former is "impersonal," the latter is "personal." At stake in the effort to define these categories is which kind of ethics should characterize professional philosophy – and what it would mean to professionalize a "personal" approach to ethics.

The impersonality of ethical theory, as Sigrist puts it, arises from what others might take to be its primary value: that its topics are those about which reasonable persons may disagree. He recites a litany of typical subjects for professional ethicists -- "abortion, torture, charity, meat eating, prostitution, organ markets, climate change, poverty, gun control, procreation, reproductive rights, and so forth" – as a prelude to his

---

[1] Schwitzgebel and Rust (2014) examine the "self-reported moral attitudes and moral behavior of 198 ethics professors, 208 non-ethicist philosophers, and 167 professors in departments other than philosophy on eight moral issues: society membership, voting, staying in touch with one's mother, vegetarianism, organ and blood donation, responsiveness to student emails, charitable giving, and honesty in responding to survey questionnaires" (293). Although ethicists held more demanding moral views regarding "organ and blood donation, charitable donation, and especially vegetarianism" (319), they were no more likely than others to pay dues, be honest on surveys, or respond to a charity incentive, and were less likely than non-philosophers to keep in contact with their mothers. Philosophers who gave two or more "suspicious" survey responses were more likely to rate survey dishonesty as morally worse, showing less "attitude-behavior consistency" than non-philosophers (311). The authors note that women were more likely to respond to the survey and less likely to be philosophers and thus that "gender could potentially have played a confounding role" (316). "It remains to be shown," they conclude, "that even a lifetime's worth of philosophical moral reflection has any influence upon one's real-world moral behavior" (320).

point: "being right" about these topics "does not make you good even when you have all the right opinions and act on them perfectly." It may be good to be right about these matters, he concedes, but right beliefs about them and right action in accord with those beliefs does not "make you good." One reason for this, Sigrist argues, is that most of these topics pertain to dilemmas few people will encounter frequently in their daily lives. I find this unconvincing: his claim that "[e]ven decisions about procreation and abortion are made at most a few times in life, if at all," might have benefitted from some workshopping with those who more frequently navigate the politics and responsibilities of birth control. More plausible, I think, is his next point: even the topics more applicable to everyday life – "climate change, vegetarianism, and poverty," in his estimation – are "mainly issues of public, not personal, concern."

Here "public" is meant in two senses: first, that what matters is not so much individual decisions as "collective patterns of individual decisions over time," and second, that the commitments arise not from the "substance of a particular person's life, but from principles that purport to hold true for anyone." The first, though it may presume a consequentialist framing on the part of the moral agent, is intuitive if "not personal" is taken to mean something like "not likely to be experienced as *my* accomplishment": there can be an alienating or dispiriting aspect to pursuing public goods for which individual decisions serve as merely drops in an ocean of action involved in accomplishing the goal.[2] The second involves understanding "not personal"

---

[2] Sigrist does not question this framing of the "public" as "not personal" or consider if it may be a symptom of, say, a specific political formation lacking in solidarity or the particular condition of an individual whose "personal matters" are rarely litigated in the public sphere rather than a general truth of the categories. However, I take it as reasonable to suppose that no matter the political formation, not every act in pursuit of a broad good like "ending poverty" will have the personal character of, say, a face-to-face act of charity.

differently, as "not taking into account *my* actually held values and commitments." What Sigrist rejects here is a formalist ethical ideal: that the purpose of morality is to produce universal principles, justified on an objective basis, that do not depend on and even override personal commitments. The general term "personal," by implication, unites these two themes: it is what arises from one's unique circumstances as relevant to ethical deliberation – purposes, values, projects and commitments – and in so arising, is necessarily experienced as "mine," as both relevant to and expressive of a singular person. The "impersonal," then, covers matters relevant to ethical deliberation that may be detached from such circumstances – including, say, beliefs that many rational persons from various backgrounds might accept – and so are not necessarily experienced in any particular way.

One need not agree with Sigrist's rejection of this formalist ideal to follow his main points: that there are impersonal and personal kinds of ethical thinking – that which "aims to be public and impersonal" and that which "arises from the substance and particularity of an individual's real life" – and that professional ethics done by and written for academic philosophers is dominated by the former. Sigrist unfortunately does not explore the possible advantages of impersonal ethical thinking. Whereas what is personal, precisely because of its unique or sensitive conditions of emergence, may not be a matter for others to properly evaluate or criticize, what is impersonal may be an accessible topic of debate for persons from widely diverse backgrounds. Put differently, debating what counts as a justified belief – it is supported by true premises, it does not involve self-contradiction, etc. – is not inherently fraught if it proceeds respectfully; debating if someone is personally committed to the right values and projects is almost

certain to be sensitive.[3] Perhaps at worst, the impersonal concerns no one – like a thought experiment with a hundred hypotheticals, a "what if?" in a void, it relates to no one's experience or values – but at best, it concerns everyone, or as many as it possibly could.

To concern everyone, I would say, is another way of defining this ideal of formalist ethics and its universal principles or procedures. For a successful system of formalist ethics that has justified both the universal applicability of its principles to moral agents and the overridingness of those principles over those agents' merely prudential or self-interested concerns, impersonality is the advantage. An impersonal ethics that is "impartial" is one in which an agent may not justify making an exception for themselves from the moral measure that they apply to everyone else just because of their claim to unique personal concerns, commitments, or feelings. For an impartial ethics, it is impermissible to take, as William James puts it, a "moral holiday" (2018, 29).

Even if one is not a strict Kantian or utilitarian, or otherwise lacks such formalist ambitions, there may be intellectual and pedagogical benefits to an emphasis in philosophical literature and research on "public and impersonal" matters.[4] Sigrist focuses instead on what may be lost in the neglect of the personal in professional ethics. He contrasts abstract questions of the morality of procreation – say, if it is "morally permissible to create another human being" – with the lived experience of his and his

---

[3] Perhaps the interlocutors will not agree on what constitutes a justified belief, on standards of discourse, or even the appropriate system of logic. But if so, it remains possible for a conversation to then proceed, perhaps awkwardly or fitfully, to debate these topics. It is possible to take standards of discourse, or logic itself, very personally. But I do not think such formal considerations are "personal" in the manner of the values, projects, and commitments that shape what strikes one as salient: such considerations are, at most, personal insofar as they reflect those values.

[4] I don't think this categorical distinction between "personal" and "impersonal" is likely to be broadly useful in pedagogy: for one, it's obviously possible to engage on a strong "personal" level with topics that, by Sigrist's definition, count as "impersonal." But it's at least provisionally helpful in showing the difference between, for example, teaching virtue ethics (1) as a set of claims on which students are tested via essays or exams or (2) by asking students to practice virtues and reflect in writing on their efforts.

wife's decisions to have children and the necessity of thinking from "the specific circumstances of unique individual lives:" for example, of questioning their own motivations, asking if they were driven by desires to please their parents or mimic their friends, or asking if that mattered. The contrast, Sigrist claims, is that the latter, personal kind of ethical thinking "cannot but alter your attitudes and behavior." It requires reflecting on, and often making a deliberate effort to transform, one's own "feelings, desires, motivations, [and] values" in a way that impersonal ethical thinking does not. What Sigrist ultimately objects to is the presumption that personal ethical thinking is "beyond the business of academic philosophy." Sigrist thus takes Schwitzgebel's concerns in a different direction. The problem, as Sigrist sees it, is not that professional ethicists are settling for a form of moral mediocrity determined by impersonal parameters – say, percentage of income donated to charity – of moral excellence; the problem is evaluating moral excellence primarily in terms of impersonal parameters. "Being good," he concludes, "is mainly about things like showing compassion to a difficult colleague, finding affection for one's spouse even when they let you down, knowing how to care less about what is less important, guarding against destructive anger, learning to forgive, and so on." It is, echoing Aristotle, "about personal wisdom," which he thinks we have not "moved beyond" but rather "forgotten."[5]

My project – this work – is a new attempt at reintroducing the personal to professional scholarship in ethics. I mean this in at least two senses. It is, first, an ethical

---

[5] Despite gestures in the direction of Aristotelian ideas like *phronesis*, Sigrist does not discuss Neo-Aristotelianism, active scholarship in Confucian virtue ethics, or the rise of feminist care ethics, all of which would fairly be called "personal" ethics by his definition. Since his concerns focus on the topics of professional ethics, however, I understand him to be arguing that even when such ethical traditions intervene in academic literature, they do so to weigh in on "impersonal" topics and so remain subordinate.

theory that emphasizes the personal over the impersonal, that begins with personal values and commitments and dissents from the formalist ideal of a morality with overriding obligations and universal applicability. But it is also an ethics that emerges self-consciously from my personal life, my unique circumstances, and seeks to be relevant to others precisely through that. I suspect that Nietzsche was right to say that all "great philosophy" has been "the personal confession of its author" (1966, 13) – and less great philosophy too. Still, in the ostensibly impersonal context of professional philosophy, what is supposed to matter is not the origin of a philosophy but the reasons given to support it. And I will be justifying my arguments by giving reasons that I think others who do not share my circumstances may accept, not merely by appealing to the experiences that have shaped me and led me to develop my claims. That should go without saying. But I want to be very clear. The personal must not serve as a shelter from critique. The point of personalizing philosophy should be to reach out to and draw in those alienated by the impersonal. It must not be an excuse to squirm away from criticism, sound and given in good faith, of the reasons one has given for one's views. I think that accountability, more so even than clarity or coherence in argument, is the essence of philosophical rigor. Or: if rigor is understood primarily in terms of accountability, and clarity and coherence in argument are valuable insofar as they facilitate accountability, then it may be possible to do philosophy that is "rigorous" because it is accountable even if it is not strictly argumentative or always clear or

coherent in the way that arguments should be. Instead of theorizing this vision of

philosophy further, I will try to put it into practice first.[6]

2. On Being for the Depressed

In titling my project "Ethics for the Depressed," I make both my intended

audience and my intended purpose explicit from the first phrase: my project is *for* a

distinct group, "the Depressed," though who falls under that distinction has yet to be

defined. That is, it is "for" this group in the sense of intending them as a primary

audience – the title written to call out to those who see the word "Depressed" and think

"That's me!" – and in the sense of being for their sake, intended to be somehow of

service to them in particular, to whatever extent that is possible. All ethical theory is for

---

[6] Rather, my project is not to theorize this vision of philosophy, so I won't pursue that further in the main text, but here I'll at least partly unpack the terms "rigor" and "accountability" as I've used them. "Rigor" evokes high and exacting standards of performance. In a professional philosophical context, "rigor" also can evoke the expectation and the pride philosophers have of and in high and exacting standards of argumentation. By loosening the association between "rigor" and "clarity," the byword of much professional (analytic) philosophy, I wish to reorient both the expectations and the pride (and social prestige which feelings of pride reflect) of professional philosophers. I use the term "accountability" here *because* the term is loaded to excess: it evokes both a narrower legalistic-ethical sense of contract-fulfillment (being literally held to account) and a broader care-ethical sense of openness or vulnerability (being figuratively held to account); its objects (accountable to whom?) may include other professionals, non-professionals specifically, other persons generally, one's own goals or personal standards, or some concept of truth or inquiry. Each of these aspects of "accountability" likely functions so differently that the term may only be useful when trying to capture them all at once, like I am here. But in a way, that's the point. To conceive of rigor as accountability would mean constantly asking oneself what counts as rigor (e.g., simplicity or complexity? insisting on coherence or entertaining incommensurable claims?) given to whom and in what sense one is "accountable" in a specific context. This would resemble the "culture of praxis" as defined by Kristie Dotson: "(1) Value placed on seeking issues and circumstances pertinent to our living, where one maintains a healthy appreciation for the differing issues that will emerge as pertinent among different populations and (2) Recognition and encouragement of multiple canons and multiple ways of understanding disciplinary validation." She continues: "What the initial valuing of 'live' issues and actual contribution ensures … is an encouragement of multiple canons and a fragmentation of justifying norms (including new and developing justifying norms)" (2013, 17). The problem for Dotson with the existing "culture of justification" is not norms of justification per se but the presumption of their being univocal and overridingly important: we assume that philosophy is legitimate insofar as it is justified, that we always know what justification looks like, and that it always matters if philosophy is legitimate. Surely such legitimacy matters sometimes, or even most times. But asking "when?" is itself legitimate.

someone: all rational moral agents, perhaps. But, likely because of its presumed universality of application, it rarely centers its audience. This ethical theory, what I call a "value ethics of engagement," is distinct from many ethical theories in three general respects. First, an ethics of engagement begins deliberation with an agent's starting ethical intuitions or what I will call "felt values," not a procedure: that is, it begins with the way in which someone is "engaged" with the world in terms of ethical significance. Second, it takes the cultivation of "engagement," the sense of ethical dilemmas or demands as personal rather than impersonal problems – problems for *you*, for *me*, for *us* – as a central theme rather than incidental to the theory's concerns or success. And third, it seeks itself to "engage" by centering its audience, by being for someone. These goals are linked, but I take the third to be foundational. The most important difference of this version of personal ethics, compared to other ethical theories, is its imagined interlocutor: that is, whom it is trying to convince.

I think that impersonal approaches to ethics tend to envision as their hypothetical interlocutor an egoist: that is, a rational but self-interested agent who is not yet convinced that moral obligations might override their prudential concerns. There are alternative interlocutors that also suit an impersonal approach of ethics – the skeptic, the subjectivist, the relativist, the amoralist (however variously these may all be understood) – insofar as these are all characters who reject some ethical considerations but in principle are susceptible to rational arguments for accepting those considerations. I take the egoist to be the most believable of these characters, one who may be opposed to the altruist in a contest that Bernard Williams takes to be the heart of ethical considerations in general.

10

This opposition is presented by Williams in "Egoism and altruism" in *Problems of the Self* (1973) and many of these interlocutors are discussed in the context of Plato's writings in *Ethics and the Limits of Philosophy* (2006). In the former, Williams does not distinguish between "ethics" and "morality" as he does in the latter, where ethics is concerned with the Socratic question of how one should live and broadly "relates to us and our actions the demands, needs, claims, desires, and, generally, the lives of other people" (12) and morality more narrowly defines ethical considerations in terms of purely voluntary acts and overriding obligations. As he later remarks, "it is a mistake of morality to try to make everything into obligations" (180). This is a mistake of depression, too.

But the egoist in particular is relevant to an ethics for the depressed, because the egoist and the depressed person are similar in an interesting respect: even if they accept reasonable standards for evaluating the moral status and obligations of others, they tend to make exceptions to those standards for themselves. The egoist makes their own ethical life unacceptably easy. The depressed person makes it excruciatingly, bafflingly hard.[7]

My project does not involve attempting to persuade a hypothetical egoist: I do not accept that moral obligations are universally overriding of prudential concerns, or even that there is a clear moral-prudential distinction; I think that committed egoists susceptible to rational argument are rare; and if I had to persuade one to behave ethically, I would settle for appealing to instrumental values like being able to maintain relationships or hold down a job. But these points are all incidental to the impersonal

---

[7] There are more nuanced understandings of ethical egoism: an egoist might not make an exception for themselves and instead be committed to the claim that nobody's moral obligations override their prudential interests. But notoriously, this view tends not to be in the egoist's interest, insofar as the egoist relies on others keeping their promises, etc. So, I operate here with the more conventional figure of a selfish person.

ethicist. The impersonal ethicist just wants the best possible justifications for their ethical positions, and the egoist stands in as a figure who would *only* be persuaded by such justifications, which, ideally, are ultimate justifications, where to deny them would mean to disavow one's own rationality. As Williams observes, this amounts to the rejection of moral luck, which is the problem of a moral agent being justifiably responsible for what is beyond their control. For example, if solely an act's consequences determine the apportionment of moral credit or blame, then unforeseeable events legitimately affect this apportionment. The act of some X who successfully donates money online to charity is justifiably called morally better than the act of some Y who attempts to do so but, due to an unprecedented computer error, finds that their money has vanished. The consequences of X's act may be morally better that those of Y's, but to say that Y's act is justifiably less praiseworthy, despite the identical intentions, due to unforeseeable events might make the moral system itself seem unfair. X was just morally lucky compared to Y.

The problem with turning to evaluating moral acts based on intentions, however, is that one's character or will may also be the product of luck. Imagine that I am a neurotic or thin-skinned person because I grew up in a family environment where I acquired the habit of being hypervigilant about the prospect of conflict or threats to secrecy or social status. I may be more likely to retaliate to perceived threats out of what I feel to be justified or proportionate defensiveness but which others are likely to perceive as unwarranted aggression. It may be fair to say that my acts arise from a traumatic psychological injury, of which I may not be aware. But I am intentionally and deliberately choosing to act "defensively," or at least rationalizing my behavior in a manner that I take to be deliberation, so it probably seems fair to hold me responsible in

some way for the impact of my actions. But am I to blame? Should my actions accrue more blame and less praise than someone who happened not to be traumatized in the past? Is that fair?

Now ethics is in trouble. Whose actions or dispositions have not been shaped by the "luck" of their upbringing? Which actions can be fairly praised or blamed and how do we know without a deep, perhaps impossibly deep, dive into every agent's psyche? Perhaps we attempt to dispense with the notion of justified praise or blame. Is there such a thing as social responsibility without praise or blame? What would it mean to hold someone responsible for their actions without justifiably praising or blaming them?

An impartial and impersonal ethics aims not just to justify being fair, but to be fair as an ethics. It aims to develop an absolute measure for how to apportion responsibility, or praise and blame, so that we may know who deserves praise or blame and for what, so that we may make sure that they get it. It aims to correct mistaken notions about who is responsible for what. No theory can make people fair. But a theory can define and justify what it means to be fair, and argumentatively challenge flawed justifications for mistaken notions about responsibility. The success of such a theory relies on its own fairness, that is, its resolution of the problem of moral luck.

I believe that the impartial and impersonal project of developing absolute or ultimate justifications for ethical positions cannot succeed. For reasons that I will explain, I assume that the elimination of moral luck is impossible. Even if I am right, that does not make developing formal theories like Kantian ethics somehow futile – it's possible to reach true conclusions from false premises, and certainly to reach interesting ones – or mean that I think philosophers should cease to pursue such projects. What concerns me is

13

what I take to most concern Sigrist: that the personal is largely neglected by professional ethicists. I hope not to afflict this neglect upon other philosophers but simply to alleviate it, and I think this requires showing what the personal has to do with the professional: that is, not merely showing that it is possible to both give good arguments and discuss one's personal values and commitments, but showing what role the latter plays in ethics.

So, my imagined interlocutor is not the egoist or any impersonal alternative. Instead, I imagine an ethics that serves a depressed person, or anyone susceptible to a condition I call "demoralization." To be demoralized is to feel that it is impossible to "be good," or to feel that one is essentially unworthy and to feel powerless to do anything about it. In more philosophically technical terms, to be demoralized is to compulsively apprehend one's own ethical status as lowered or inadequate such that no ethical act can seem to restore it, thus producing a sense of inhibited ethical agency or loss of motivation. When demoralized, it may be possible to imagine that other people are or can be good, that is, ethically worthy or fulfilling ethical obligations to an acceptable degree. But the criteria that apply to others do not apply to the demoralized person: perhaps they think that they have gained something that sets them apart – a special knowledge of the impossibility of their being good – or they think that they have lost something that others do not even know they have, or they think that both are true at once. What the demoralized person needs is not to be persuaded that moral obligations are overriding of their prudential interests: if anything matters to this person, it is moral obligations; they may feel lucky to have any interests left. What the demoralized person needs, I think, is to be persuaded that it is possible to "go on:" that some acceptable form of ethical status is attainable for them, that they may live as they should, that ethics – meeting ethical

obligations, fulfilling ethical values, ethical life – is possible for them at all.[8] Before I

elaborate on the distinction between depression and demoralization, I will say more about

this "sense" of one's own ethical status and why it should serve as a starting point for an

ethical theory.

In describing the "existential guilt" of depressed persons, the feeling of being

deeply and irrevocably guilty for no specific reason and before no specific person,

Matthew Ratcliffe emphasizes "a focus on past deeds, recognition that effects of one's

deeds are unchangeable, an awareness of estrangement from others, a sense of having

done wrong or of being intrinsically flawed, and an anticipation of being harmed or

punished" (2015, 134). Such experiences of existential guilt "cause considerable

suffering and are sometimes singled out as the most troubling symptom" of depression

(135). In having no clear reason or object, existential guilt also lacks an important futural

dimension of conventional guilt, the prospect of eventual forgiveness, and so may

preclude, or seem to preclude, hope.[9] Knowing this, we may work backwards to what it

---

[8] Though there are many useful parallels to demoralization in other scholarship, I take one especially strong parallel to be "moral distress" (Fourie 2017) in nursing ethics. Ed Yong (2021) describes it as "the anguish of being unable to take the course of action that you know is right." He writes in the context of the COVID-19 pandemic: "Health-care workers aren't quitting because they can't handle their jobs. They're quitting because they can't handle being unable to do their jobs. Even before COVID-19, many of them struggled to bridge the gap between the noble ideals of their profession and the realities of its business. The pandemic simply pushed them past the limits of that compromise."

[9] Fyodor Dostoevsky's novel *The Brothers Karamazov* (2002) expresses a similar theme: "each of us is guilty in everything before everyone, and I most of all … verily each of us is guilty before everyone, for everyone and everything" (289). I take existential guilt to lead compulsively to a series of errors: first, a mistaken sense of one's being ethically inadequate or unworthy; then mistaken conventional guilt, e.g., "I must be wronging X in Y way," as though this sense of unworthiness, in being subsumed under the form of guilt, searches out in memory and so mistakes its own cause; then, since guilt which arises from no wrong cannot be redressed, a mistaken sense of powerlessness to alleviate this sense of unworthiness. It is possible that existential guilt might reveal a truth, that we really are guilty before everyone for everything. I think it is best to discount existential guilt entirely as a source of truths, given its propensity to produce misattributions of its own cause, i.e., mistaken senses of conventional guilt. But such an idea is reminiscent of (though not equivalent to) the Christian concept of original sin and, separately, Levinas's ethics of alterity. A key difference is that existential guilt is not alleviated by any redemptive act of willing, whereas original sin might be redeemed through faith. I will further discuss Levinas in Chapters 5 and 6.

would mean to the depressed person to at last "be good": it would be whatever amounts to the alleviation of their existential guilt.

For an ethical theorist, this may seem like a bad place to start. The depressed person is not guilty of anything. Their experience *has the form* of guilt, minus its hopeful future. "Being good," so conceived, is defined by an involuntary or compulsive negative self-evaluation that seems severed from any typical conception of moral status. Also, if it is likely that depressed persons have mistaken intuitions about their ethical status due to an excessive sense of their ethical obligations, it may seem perverse to construct an ethics that defines their *actual* obligations, and thus what will count for that ethical theory as "mistaken," by beginning with those intuitions. (If we end up agreeing with a depressed person that they really are unworthy, we've gone very wrong.) Even so, I want to start here, and not just because I'm depressed. Michael Cholbi argues that a depressed person's indifference or "listlessness" is not an ethical hazard because depressed people are more, not less, likely to be concerned with others' well-being and with their own moral adequacy (2011, 38). On his reading, depressed persons become numb to their own prudential interests, not their ethical considerations or obligations, with which they might even be excessively concerned. Even if an ethics for the depressed will not sustain a strict distinction between the prudential and the ethical, I take Cholbi's point: depressed persons perceive themselves, not others, as unworthy. Even if they mistake existential guilt for the conventional, they look outwards for what feels like forgiveness.

All this is to say: I predict that beginning an ethical theory with the question "What will alleviate a depressed person's existential guilt?" will lead to a prescriptive answer that will also satisfy questions like "What is ethically good and why?" even for a

16

person who is not depressed; but beginning with the latter question will not satisfy the former and, for a depressed person, may exacerbate the pain associated with the absence of that answer. And I predict that a depressed person does not simply retain most ethical considerations despite existential guilt; instead, what will be generally recognizable as ethical considerations may follow from existential guilt and the pursuit of its alleviation.

These predictions anchor my response to a strong fundamental objection to my project: that this interlocutor, depressed or demoralized or whatever, needs therapy, not philosophy. "The depressed person is compulsively mistaken about their own ethical status in a manner resistant to rational argument," this objection goes, "and philosophy is characterized by rational argument. This makes for a poor fit. Even if philosophy may be therapeutic, centering an ethical theory on achieving a therapeutic goal will at best set up an uphill battle and at worst twist philosophy into something unrecognizable to philosophers. If a depressed person wants a philosophical theory to solve a therapeutic problem – for example, to justify their own ethical status as adequate and so 'correct' their compulsive mistake – rather than go along with them, we should say that their problem will not likely be responsive to theory and that they are thus sadly mistaken about what will actually help." I agree with this objection that philosophy, and ethical theory as a subspecies, is and should be characterized by argument, and that compulsive thinking or behavior is generally unresponsive to argumentative reasoning. But I ask that those who support this objection follow me long enough to see if my two predictions hold up. My depressed interlocutor is, I think, driven to seek the consolation of philosophy. By assuming that philosophy has something to offer them, without pretending to substitute

17

philosophy for therapy, we will see what possibilities philosophy has neglected thus far, to the disappointment of the depressed.

Thus, while I follow Bernard Williams in my understanding of ethics, I diverge from him but in my goals for ethical theory. This is best illustrated by considering Williams's brief discussion of ethical conviction in *Ethics and the Limits of Philosophy*. Based on his claim that reflection can destroy ethical knowledge by undermining "thick" concepts (which have both evaluative and factual dimensions), Williams understands ethical conviction not in terms of knowledge or certainty but "confidence," the capacity of one to both maintain and question one's ethical attitudes and concepts, which he regards as more dependent on social context than argumentation. But argument has its place: "Social states can be affected, one way or another, by rational argument. Moreover, if we try to generate confidence without rational argument or by suppressing it, we are quite likely to fail, but, besides that, we shall be sacrificing other goods" (170). Considering his doubt that "ethical thought is made more rational by deploying ethical theory" (xii), I take Williams to be saying that ethical arguments are unnecessary and insufficient for instilling ethical confidence but still significant in avoiding the weaknesses of dogmatism. This contrasts with moralists who may, consistent with their views on volition and obligation, presume mistakenly that conviction will arise from argument, which is to say that theoretical justification will contribute to practical resolve.

I think such moralism, what I call "the strictness of the depressed," is tempting to a depressed person who turns to philosophy in search of answers. And I am attempting to develop an ethical theory that takes instilling "confidence" as its highest goal. But I believe I follow Williams in thinking that we can get from ethical considerations broadly,

18

in his sense, to a better sort of ethical theory than what he calls "morality." What I have done that Williams has not is take the figure of the depressed person seriously as a philosophical interlocutor. When in *Morality* (2015) Williams considers the question "Why should I do anything?" as someone's "expression of despair or hopelessness," he rightly says that it is unlikely that "we could *argue* him into caring about something … What he needs is help, or hope, not reasonings" (3). He is also right to say that it is not "a defeat for reason or rationality that it had no power against this man's state" (4). But this question, even when it is primarily an expression of hopelessness, is not *just* that: it takes the form of a request for reasons.

I think that Williams interprets the depressed inquirer as only appearing to seek reasons or argument but as actually just seeking exhortation or encouragement. But there is a peculiar way that the depressed inquirer seeks a blend of both, a blend for which my vision of ethical theory aims. Williams's fundamental oversight here is to presume that persons are either autonomous, as in responsive to reasons, or not. But there is a middle ground between those who are autonomous and those who are not. Depressed persons inquire from that middle ground. They are who I call the "aspiringly autonomous": they are responsive to at least some reasons and, more importantly, they *want* to be responsive to more reasons. When faced with such an interlocutor, it would be a mistake for philosophy to either proceed as usual or turn away. The question becomes what a correlate "middle ground" for philosophy would be.

Having said that, I return to the distinction between "depressed" and "demoralized." I take demoralization to be a consequence of the existential guilt that Ratcliffe finds characteristic of depression: because existential guilt, unlike conventional

19

guilt, does not involve a specific wrong committed against anyone in particular, the existentially guilty person regards themselves as unworthy no matter their actual ethical status and as essentially so because no forgiveness is possible.[10] I write of an "ethics for the depressed" and not an "ethics for the demoralized" because I think "depressed" is a term that people more commonly ascribe to themselves, whether or not they have been medically diagnosed with some form of depressive disorder, and because what I am calling demoralization resembles symptoms of depression. In the DSM-5 (American Psychiatric Association 2013), the diagnostic features of major depressive disorder are "either depressed mood or the loss of interest or pleasure in nearly all activities" and may include "changes in appetite or weight, sleep, and psychomotor activity; decreased energy; feelings of worthlessness or guilt; difficulty thinking, concentrating, or making decisions; or recurrent thoughts of death" (163). An episode is accompanied "by clinically significant distress or impairment in social, occupational, or other important areas of functioning."

I think that those who think of themselves as depressed are likely to recognize themselves in the demoralized person and that they will find the philosophical account to follow especially resonant. But I do not mean to presume in advance a medicalized context for demoralization or the inclusion of persons diagnosed with depressive disorders who do not experience demoralization as I have defined it. The DSM-5, in contrast to the DSM-IV, separates "Depressive Disorders" from "Bipolar and Related Disorders," explaining: "The common feature of all these disorders is the presence of sad,

---

[10] Demoralization might arise from feelings or dispositions other than existential guilt, like lowered self-esteem. In this project, I will focus on developing an account of demoralization as it arises from existential guilt to retain a consistent focus on a depressed interlocutor, but I do not argue that it necessarily so arises.

empty, or irritable mood, accompanied by somatic and cognitive changes that significantly affect the individual's capacity to function. What differs among them are issues of duration, timing, or presumed etiology" (155). Like Ratcliffe in *Experiences of Depression* (2015), I do not presume to speak for those diagnosed with Bipolar I or other manic-depressive disorders. This does still raise the question if someone with a manic-depressive disorder would count as a demoralized person by my definition, and thus if an "ethics of engagement" would apply or be relevant to them. Someone experiencing a manic episode likely would not benefit from being or feeling more "engaged," if that means "elated." I aim to attend carefully to the question of the scope of an ethical theory like mine, that is, whom it obligates, what it is likely to obligate, and on what basis.[11]

Similarly, not all those who count as demoralized persons may be depressed if one distinguishes depression as an "internal" disposition from "external" social or material features of the world. This of course depends on one's definition of depression. But I think it is plausible that someone who, say, otherwise experiences themselves as personally stable and loved or lovable, rather than as being personally unstable and unloved or unlovable, may still experience the world as deeply hostile. For example, a member of a stigmatized or marginalized group may face unjust treatment from, and the withholding of social and material resources by, members of a higher status group. This

---

[11] In Chapter 5, I distinguish between "elation" as "a cluster concept referring to the pleasant sense of excitement that can accompany activity that seems worthwhile or significant in the moment, including senses of being present or in the zone" and "engagement" as involving elation plus a "sense of stability," which lack the sense of precarity or isolation that may make experiences of elation seem brittle or fragile. The concept of engagement may remain relevant in the context of manic-depressive disorders insofar as it is distinguished from simple elation. I also believe that the moral psychological and metaethical claims about motivation and value, and regarding the viability of a value ethics grounded in attention to felt values in the present, that underpin the ethics of engagement would be relevant to an interlocutor that must cope with an oscillation between depressive and manic episodes. However, what it would mean to enhance this interlocutor's sense of agency might greatly differ, not just in terms of specific normative requirements but in the account of ethical deliberation and its emphasis on attentiveness to the present.

person may feel that it is impossible to "be good" not because they internally assess themselves as unworthy but because they understand or conclude that they will be seen as unworthy by members of the higher status group, no matter how much they personally change or accomplish, and that it is how this higher status group sees them that determines their worth. The fact that I begin this ethical theory with existential guilt may limit its application to persons demoralized by systemic injustice. One might even argue that a political theory that responds to systemic injustice, rather than an ethical theory, is the more appropriate response to demoralization. But if I am right to later conclude that existential guilt, in its manifestation as what I call the "call of the other," is one way of revealing a more general condition of vulnerability, then an ethics for the depressed may become an ethics for the demoralized. Still, constructing an ethical theory for the depressed is my primary goal.

When I use the term "theory" or "philosophy," I continue to think of it largely in a "professional" sense, just as Sigrist does: I am interested in that word's ambiguity. Sigrist does not define "professional" outright but seems to mean straightforwardly "what is done by professional ethicists:" that is, the typical content and style of academic research and instruction. Ideally, the professional is linked to the rigorous, to high and exacting standards of performance. For professional philosophers, that performance is usually argument or definition. Even in an ideal case, rigor tends to also mean jargon, esoteric shorthand that trades accessibility for efficiency. But, setting aside the ideal, to be a "professional" is also a status symbol, and academia, being competitive and hierarchical, is rife with markers of status. To be professional is to command some authority. To do something unorthodox like an "ethics for the depressed" in a "professional" way is to

attempt to confer authority on that project. One reason not to simply say that I want to be "rigorous," then, is that I am interested in both authority and accountability. I want to elevate the status of the demoralized person as an interlocutor for ethical theory; I want to be held accountable as an ethical theorist; I believe that I must expand the concept of "rigor" in professional ethics beyond sound argument to do the former but I must not give up the latter. I want to lend the "seriousness" of the professional to a project that exhorts the demoralized, therapeutic though that goal may be, and to be justified in doing so. I want professionals to think of rigor primarily in terms of accountability, and not only to their peers. And, of course, I am interested in the cliché: that a professional is one who professes, who gives full expression to a way of being, who practices what they preach.[12]

I said that I would explain why I assume that the theoretical elimination of moral luck is impossible. I do not know for sure that it is impossible, though I currently judge that there is no good reason to think that it is possible. My real reason is not purely intellectual but at least partly sentimental. I think that depressed persons are compelled to seek an absolute measure of moral worth as a form of absolution: if one exists, then it becomes possible to know when one has achieved moral worth without relying on one's own faulty judgment. Impartial ethics promises such a measure and then, to my mind,

---

[12] This move, the critique of the orthodox self-conception of philosophy as part of a turn toward a more personal way of philosophizing, resembles Pierre Hadot's (1995) concept of philosophy as *bios* or way of life, which takes Western philosophy to have largely spurned its ancient origins as a spiritual practice of self-transformation, with argument and theory being means to that end. Insofar as I take argument to be one component of a larger exhortative project, all of which I take to be rightly called philosophical, and I take professional philosophy to generally and unnecessarily be restricted by its practitioners to argument, I share this view. However, I would distinguish "transformative" and "therapeutic" goals for philosophy, and I am more comfortable with the latter than the former. Briefly: I think it may be possible to do ethical theory well in a way that both argues well for what should be viewed as good and encourages the audience to see themselves as possibly good, or that helps them to cope if not to change more radically. But I am not sure how I would make a contemporary (non-ancient) ethical theory align with broader spiritual or existential practices that help to instill virtue in any further respect, and I would be politely skeptical of efforts to try.

does not deliver. Every theory has unresolved problems. But to me, the matter of an absolute measure of moral worth is not just a theoretical problem. Its prospect can never just be a curiosity to the depressed. It will be experienced as imperative. It is cruel to, even unintentionally, dangle the prospect before them, ever out of reach. There must be a better way.

3. Ethics and Exhortation

A professional personal ethics for the demoralized person – or, to borrow a phrase from William James, "hortatory ethics" (1983, 130), an ethics of exhortation – involves (1) explanation, (2) argument or justification, and (3) persuasion or exhortation without dissolving the distinctions between them. I understand explanation as descriptive: for example, describing the rain cycle in order to explain why it rains. I understand argument or justification as normative: for example, putting forth reasons why one should believe a textbook that explains the rain cycle. And I understand persuasion or exhortation as motivational: for example, rhetoric aimed at inspiring others to become meteorologists. The differences are important: a good explanation is not necessarily a justification for anything – just because something happened does not mean that it should have happened – and a justification should not need to be inspiring to count as acceptable. To argue that professional ethics may be accountable for attempting all three is not to blur the lines between these tasks. Rather, when one considers the challenges faced by the demoralized person, who may believe wholeheartedly that moral obligations override their interests and desperately wish to be a good person yet feel that it is impossible for them as individuals, these tasks become linked in specific ways.

24

For one, explanation can be exhortative. In "Autonomy and Ethical Treatment in Depression" (2010), Paul Biegler argues that "the self-knowledge acquired through psychotherapy assists the person with depression to make more autonomous decisions in relation to the object, or trigger, of the depressed response" (179). Biegler, following the Appraisal Theory of psychologist Richard Lazarus, takes emotions to contain or precede "evaluative judgments about whether given contingencies are likely to benefit or hamper the individual's important concerns" (182). If emotions and interest-relevant (or "material") factual information combine to "guide behavior aimed at preserving the significant interests at stake," then "greater autonomy will likely ensue in cases where accurate emotional evaluations are viewed as such by the individual, but also in cases where dubious evaluations are treated with a requisite degree of suspicion." Under depression, however, "emotions give rise to false appraisals with much greater frequency than normal," and even when depression offers an appraisal function – as a response to onerous life circumstances, for example – "perceptual biases make it difficult for the person with depression to glean that information." For example, the person with depression "tends to overestimate her role in generating the adverse result, and underestimate her capacity to rectify it" (183). Depression is not only distressing but also unconstructive. But what makes it unconstructive, for Biegler, is not just the false perceptions it involves but a failure to understand how its perceptual biases operate. He cites studies that conclude "up to one-third of those with depression believe it to be a primary disorder of brain chemistry that is independent of the effects of external events" (184), thus obscuring what possibilities might exist for acting to counter depressive perceptual biases. A proper explanation of how depressive episodes might be triggered

thus can have an exhortative effect: it not only reveals possibilities for action that might meaningfully change one's circumstances but offers them to someone likely struggling with an oppressive sense of the impossibility of such change. But the exhortative effect of explanation depends on an attentiveness to context, particularly one's audience: correct descriptions are not guaranteed to inspire action.[13]

For another, practices of justification implicitly rely on some basic level of motivation to give and accept reasons. This otherwise trivial point – any and all practices presumably rely on some motivation to participate – becomes significant in the context of demoralization. The demoralized person is not likely to be less interested in justification. But two connected factors may change: the emotional character of the interest in justification and the individual's standards of acceptability. The interest in justifying a course of action may turn desperate: the demoralized person may come to think that having a "good enough" reason or purpose behind acting could lead to overcoming the inhibiting sense of the impossibility of acting in a way that could modify their ethical status, that is, counter their perceived unworthiness. Accompanying this desperation may be more demanding standards of justification, especially when others are willing to take values in question to be intrinsic. For example, when recommending a course of physical exercise, claiming that it will be good for one's health would normally be taken as a decent justification for the activity. But for the demoralized person, this may only raise uncomfortable questions as to the value and purpose of health in general or, more likely, their own health specifically. This very quality of self-exception, in which the

---

[13] An overconfident person with an exaggerated sense of their own power or agency would ideally be made more circumspect, not be inspired to further action, by a true account of their capacities.

demoralized person can take a general reason to act as good for others but not good for

themselves, may intensify their sense of isolation. The motivation to participate in

practices of justification, then, comes not from proving to a demoralized person that they

are just like any other person – or that their self-exception is irrational – but by

welcoming the demoralized person by addressing their concerns and perspective in both

substance and style of discourse.

The personal-professional style I am attempting to develop here serves two

purposes. One purpose is, as discussed, meeting the complex challenge of properly

addressing what I have called a demoralized interlocutor. The second may seem more

mundane but is worth noting: some evidence for my claims regarding motivation and

value will rely on my own phenomenological reflection about the structural elements of

experiences I have had. The quality of that evidence may depend to some degree on the

reliability of my testimony about my own experience. It is typically in the interest of

persons giving such testimony to appear "professional," that is, credible.[14] But there is

also a matter of philosophical accountability at stake: when it is *necessary* to explicitly

give testimony as evidence for a philosophical position, it is possible to draw a line

between when reasons are being given that anyone in principle should be able to accept

and when an account is being offered that must, to some degree, be trusted. I will be, for

---

[14] The fraught history of credibility in philosophy – its racist, sexist, classist, and transphobic connotations – is implicated in what it means to appear "professional" as a philosopher, especially when discussing "personal" matters, as much as it would be nice if "professional" is or could just be truthfully identified with "rigorous." There is also a separate question of when testimony is required, that is, what set of starting assumptions is taken for granted. In advocating for taking the perspective of the demoralized person as a starting point, I must work to render such a person plausible to my reader; that I find the amoralist basically implausible has not, so far, motivated anyone to try to convince me that they exist and are persuadable. But this is because the amoralist stands in as a universal interlocutor, a litmus test for justification in general, and so need not even be rendered explicit *as* an interlocutor; the demoralized person is particularized.

example, making phenomenological claims – claims about structural elements of experience that I evidence by reflecting on claims about my *own* experience – to show the inadequacy of preexisting concepts in moral psychology. In that case, where I believe I have no recourse *but* testimony, for my argument to succeed it is not enough to trust in my good faith or examine my reasoning: at the very least, one must find it plausible that someone could have such experiences. A personal-professional style involves being clear and specific about when trust is being asked for, to what degree, and to what end.

In the content and style of my writing, I hope that I have successfully transcended the scholarly divide between analytic and continental philosophy and so enacted a philosophical pluralism that clarifies and solves an otherwise obscure and stubborn problem. What I have thought I could make beautiful, I have tried to make beautiful, and what I have thought I could not, I have tried to at least make clear. I hope also that I have done all this in a way that does not call attention to itself and seems natural, even effortless. It was not effortless. I worked very hard and, sometimes, I felt very alone. If in places I seem embittered, it may be because I reel at the senselessness of a world in which this project seems to me to be made so difficult to complete.[15]

What will follow in my account of an ethics of engagement are many explanations and arguments meant to be taken on their own merits, regardless of the trustworthiness of the one putting them forward. But the success of this project, to my mind, does not depend on their success: it depends on just one reader taking just one more step. It would be better in the long term if that step were in the right direction, so to

---

[15] One of the best accounts of this senselessness that I have read is Kevin Birmingham's 2017 Truman Capote Award acceptance speech, "The Great Shame of Our Profession."

speak, and if I were to be right about what that right direction was. The point is not to broadly prize action over truth, which to me smacks of a fascistic sort of nihilism, but to frame rigor, as a set of habits which tend toward truth, as ultimately in service of improved exhortations to act.[16] I would like to encourage more philosophers to think in this way, and to think of thinking in this way as properly philosophical.

If that were to happen, I think I would feel more at home in philosophy. I want to write philosophy in the style that I have always wanted to read it, rigorous and resonant, written for someone like me. I have come this far in pursuing philosophy while feeling almost constantly that most of it has been written for someone else: someone who either takes for granted what I most question – that what counts as a good reason for others to act counts a good reason for me – or seems to misunderstand this question as coming from the egoist's place of self-interested self-exception, or something along those lines, and thus merely as a problem to be overcome. I do think of demoralization as a problem, but it is not only a problem: for me, it is the origin and the motive of any philosophical reflection whatsoever, so urgent and all-consuming that it subverts or annihilates intellectual curiosity. I have learned to admire philosophy that comes from a different place than mine, but that admiration is no antidote to my loneliness. The project of writing this is an exhortation to myself – an insistence that it is possible, in general and for me, for philosophical rigor to meaningfully serve in personal struggles – and it cannot succeed at that unless I am able to believe it may also serve as an exhortation to others. I struggle to imagine an egoist so stubborn or relentless as the questions that possess me.

---

[16] Musashi Miyamoto writes in the Book of the Void, the final chapter of *The Book of Five Rings* (1974): "In the void is virtue, and no evil" (95). I creatively misinterpret this phrase: for a deeply demoralized person, who takes neither beliefs or desires as reasons or motives to act, to act purposefully is to affirm life.

Why exist if not to be good? And why should being good, even if it is possible for others, be possible for me?

4. Biographical Note

Nora Hämäläinen begins *Descriptive Ethics* (2016) with refreshing bluntness: "Academic writing is, with good reason, governed by an ideal of impersonality. We report the end product, the scholarship and polished reasoning. The authors' struggles are not interesting" (ix). I love these lines. In what follows, I ignore them.

My mother, who separated from my father before I was born, was an academic and an alcoholic. Children of addicts tend to develop an exaggerated sense of personal responsibility, blaming themselves for their parents' problems, and in that regard, I'm no different.[17] But my mother's brilliance, her towering intellect and her struggle for its recognition, and her expectations for me lent my sense of familial obligation a peculiar valence: the path to solving our problems led through study. I believed that I could save her just by being smart. This wasn't my invention. I would dread being sat down for the regular recitations of my elementary school report cards, where every "Meets Expectations" became, in her words, a "C" and every "Needs Improvement" became an "F," when she would repeat to me that if my grades got worse, we would – somehow – lose our house. I accepted this readily, and not only because I was impressionable. I

---

[17] In "Risk Factors Among Adult Children of Alcoholics" (2007), Hall and Webster summarize the research on this topic: "Responsibility in the alcoholic household becomes a central issue … Children growing up in this environment may take on a sense of hyper-responsibility where they believe they are responsible for anything bad that happens in their family" (497). Assuming the role of caretaker for the family "is very attractive because it creates the illusion of feeling powerful and being important and needed. The opposite side of the coin, however, is that there is very little or no control over events. The child feels responsible for the behavior of the alcoholic parent but does not have any real influence or power to change the behaviors of the parent or the family dynamics."

believed that I was a member of an elect. What was impossible for others was, should, must be possible for me, because I was special, because I am her son.

I was defined by a singular purpose: to save one person from herself. And so I was defined, really, by the struggle to cope with the impossibility of fulfilling my singular purpose. A person subject to this persistent stress – this trauma – experiences a foreshortening of the future: only the present, painful or pleasant, threatening or safe, feels real.[18] Growing up, I thought about leaving home or my future only fleetingly and in the vaguest terms: a life not defined by my mother – my mission – seemed unimaginable. Even now, I cannot think of my future self as another person whom I might benefit, whose path I might clear for him, by present prudent acts: as soon as I try to imagine him, he feels irrelevant to me, a stranger without a face. Self-motivation becomes a cruel and constant dilemma when your future self is someone else. The only way through that I've found so far is a more careful attention to my present self, to the subtle tenor of my drives and motives and what I can gauge of their underlying conditions within and without me, as though my being were a throughline of feeling. All this has led me to believe that it is possible to dissociate even from desire, that it is not enough just to want and to believe that what one wants is good, that the very possibility of motivation cuts deeper than that.

---

[18] Ratcliffe, Ruddell, and Smith (2014) refer to the "sense of foreshortened future" defined as a symptom of post-traumatic stress disorder in the DSM-IV (American Psychiatric Association 2000, 468). The authors understand this foreshortening not as a "cluster of interrelated judgments" about the content of what the future holds – career, family – but rather a more fundamental change in "the experience of time," an "altered sense of temporal passage." They write: "When the person looks ahead, the future lacks structure; it is not ordered in terms of meaningful projects, and so a coherent sense of long-term duration is absent. Hence the all-enveloping dread she feels before some inchoate threat is not situated in relation to a wider pattern of meaningful temporal events There is nothing meaningful between now and its actualization, and so it seems imminent."

I remember the day that I, to my mind, became a person. When I was 16, I chose to leave my mother. She had been compelled to seek treatment – graduate students had complained – and then left it against doctors' advice: I had lived with my father for a year but returned to her, where I felt I was needed. But that feeling had begun to ebb and leave a hollow. A DUI verdict required her to attend Alcoholics Anonymous meetings; I discovered that she would go out drinking during the allotted times. The deeper her lies, the clearer my powerlessness became. That day, leaving home for an aimless walk, I met her out front returning with wine. The strongest person I knew wouldn't meet my gaze. I understood, finally, that nothing I did would make a difference. Yet I felt that if I allowed her to walk past me without acknowledgment, I would entirely disappear. I stepped in front of her and reached into the bag. I wrested a bottle from her – carefully, not wanting to hurt her – and she shouted for help as though I were a stranger. I walked to the storm drain and knelt to crack the glass – carefully, not wanting to make a mess – against the curb, watching the wine flow into the darkness. When I looked up, she was already walking away, back to get more. I left the next day. She died by suicide.

How am I supposed to live? I must choose a purpose for myself. I cannot recognize a life without a chosen purpose as a life: it might be possible for others, but not for me. So, I do not need to be convinced that I must choose a purpose. Honestly, I'm not sure it's actually true that I must, or that there's a proper justification for the claim even if it is true, but it seems that regardless, it's the only way I can go on. Let's say that it's true, not just necessary. Let's say that this compulsion to choose a reason to go on is also the obligation of an examined life. What, then, would it mean for my choice in purpose to be

32

justified? (If I found all purposes to be equally good or justifiable, presumably I wouldn't have a problem.)[19]

For someone like me, there emerges a distinction between being justified and "feeling justified" and, perhaps oddly, the latter presents itself as the greater problem. Being justified is, more or less, a matter of giving acceptable reasons: it involves what others will (or should) take to be sufficient to accept a claim as a reason. Feeling justified is categorically different: it involves what *I* will (or feel that I should) find to be sufficient to accept a claim that I *give to myself* as a reason. For any good account of being justified, feeling justified is irrelevant. A feeling of reluctance to accept a claim one is considering could easily be the product of irrational prejudice.[20] But this notion of "feeling justified" still has something important to show us, even if it does not – and should not – serve as a ground for justification writ large.

What makes "feeling justified" resemble justification at all is that it involves permission to go on. There are, I expect, more differences than similarities between "permission" given by others in conversation and "permission" granted to oneself by oneself, especially when the latter permission is of a sort that seemingly cannot be given by a conscious act of willing.[21] But my fundamental problem isn't that others won't find my beliefs or desires acceptable. It's that beliefs and desires I find "acceptable" one day – that I find to be sufficient reason and motive to act, that are good enough for going on –

---

[19] Or maybe I'd have a much bigger one.

[20] This simple example belies a deeper issue epitomized by the Sellarsian critique of the myth of the given: any appeal to intuition in justifying a claim to knowledge (like, a gut feeling) must present that intuition under the rubric of some concepts in the process of justification (like, words to talk about the gut feeling).

[21] I am distinguishing here between a volitional act, like choosing to break a rule, and an experience of the limits of one's volition, as when a demoralized person finds that an act seems impossible for them yet possible for others.

may abruptly cease to be "acceptable" the next. The seeming impossibility of my future and the immensity of my guilt may place me outside the realm of the reasonable – of the set of conventions that help to determine what is generally to be taken as a good enough reason – yet it is that very placement that compels me to ask for reasons. I'm stuck.

What can ethics do for someone like me? It can help to alleviate the crushing burden of the imperative to be good by constructing, consistently and coherently, an acceptable and recognizable approach to ethics that incorporates personal projects and commitments, one that does not rely on its obligations universally overriding personal feelings and affinities. It can reveal the temporal underpinnings of the motivational structure of belief and desire, their implicit reliance on a felt and initially inchoate sense of possibility, rather than taking the presence or absence of desire to be foundational to motivation. It can envision an agent embracing present feeling, beginning ethical deliberation with the value paradigms – fairness, care, loyalty, authority, purity – that tend to structure ethical attention and framing norms as predictions as to what will extend these values. It can develop a practice of intention formation in ethical deliberation and, concomitantly, practices to counter the dissociative qualities of demoralization. It can, in all these ways, show a demoralized person that it is possible to be good.

You can't just theorize your way out of depression. But explanation can be exhortative. Depending on how you describe the world, different things seem possible. And justification depends on motivation. People like me, I think, search desperately for good reasons to go on. We want to believe that a future is possible for us, but we can't. Yet that "can't" is tricky. What feels impossible isn't necessarily so. Maybe the next

explanation, the next justification, the next exhortation will make a little difference. Maybe not. But the chances may be higher – may be – if we're addressed directly.

I'm going to try.

5. Summary

From an ethical standpoint, existential guilt amounts to a two-part compulsion: 1) the compulsive assertion or sense of an absolute and vague threat that disrupts the formation of intentions and actions, and 2) the compulsive taking of this causal disruption as a reason for inaction. Correspondingly, I present my dissertation in two parts.

The first part of my dissertation, comprising Chapters 2 through 4, largely concerns the first aspect of existential guilt: it is a metaethics for the depressed, or "ethics as a reliable guide" as a response to "demoralization" and "hypermoralized deliberation." There I challenge what I call the Stocker-Smith account of depressive loss of motivation as being a loss of desires and argue instead that it involves the defeating presence of what the phenomenologist Matthew Ratcliffe calls "pre-intentional" mental states, a category that I redefine and expand to include second-order "quasi-beliefs" and habits of feeling, that interfere with intention formation and action despite the persistence of desire.

The second part of my dissertation, comprising Chapters 5 and 6, largely concerns the second aspect of existential guilt: it is a normative ethics for the depressed, or a "value ethics of engagement" premised on "contingent value ranking." After demonstrating in the first part that depressed persons may retain their desires in depression, I premise a value ethics upon what I call the consistent desire for a "sense of stability" in response to experiences of precarity and isolation. From the phenomenology

of value, I develop a concept of the heart as the set of "felt values" or intuitive value paradigms that are themselves pre-intentional states or dispositions.

I thus attempt to structure a complete ethical theory, integrating plural philosophical traditions and founded on the phenomenological category of pre-intentional mental states, in response to the presence of existential guilt as experienced by an otherwise reasonable interlocutor. I put an orthodox style of philosophy in service of an unorthodox agent: one who is "aspiringly autonomous."

My primary contribution to metaethics and moral psychology is my adaptation of Ratcliffe's concept of the pre-intentional from phenomenology. I redefine the pre-intentional as a category that encompasses both 1) intentional states that are directed toward other intentional states, rather than directly toward objects in the world, and 2) seemingly non-intentional or intentionally ambiguous states that affect other intentional states.

In Chapter 2, "Depression and Motivation," I discuss an example of the first subtype: a quasi-belief that responds to and undermines a belief or desire. For example, a depressed person might desire to solve a problem and then, in the process of deliberating or acting on that desire, be confronted with a quasi-belief like "this desire is impossible to fulfill because I am unworthy and incapable." The second-order quasi-belief is directed toward the first-order desire, which is an intentional state directed at some object in the world. This turns deliberation into rumination or spiraling. The desire continues to exist, but it is undermined by the quasi-belief in its generating of plans or action, and the quasi-belief is not responsive to evidence. In fact, the depressed person probably believes that they *are* capable, a belief that contradicts the quasi-belief. Insofar as the quasi-belief

36

persists despite refutation, that contradiction is irresolvable. Both due to the quasi-belief itself and the compulsively irresolvable contradiction between the belief and quasi-belief, the depressed person is stuck, unable to proceed in planning or acting.

In Chapter 3, "The Pre-Intentional," I discuss an example of the second subtype: is a habit of feeling threatened. In this interpretation, the relevant mental state to a depressed person's loss of motivation is not another intentional state (a quasi-belief directed at a desire) but a non-intentional or ambiguously intentional state like a pervasive and vague feeling of threat from everywhere and nowhere. What then happens is that the depressed person reacts to this feeling first with a fight-flight-or-freeze response and then by attempting to resolve the threat, say, by searching for its origin, which risks going on indefinitely because the depressed person is compulsively unable to experience a sense of resolution.

I include both quasi-beliefs and habits in the category of the pre-intentional to emphasize that depressive loss of motivation is introspectively opaque. I do not know if depressive loss of motivation, or demoralization, is caused by 1) quasi-beliefs that contradict beliefs about the possibility of success at one's goals or 2) dispositions of feeling threatened that then compel automatic defensive behavior or inferences that threats exist. Either is a plausible account. I do not think that it is possible to tell which account is more accurate either by looking inward or to the consequences, since both sorts of pre-intentional states function equivalently. But I argue that it does not matter which is the more accurate account of depressive loss of motivation if the category of the pre-intentional is defined to include both. What matters is that I can show that when depressed persons become demoralized, they are not, for example, necessarily losing

37

their preexisting desires or values. If depressed persons may plausibly retain those even in depression, I can argue for an ethics that begins with them.

In Chapter 4, "Depression, Deliberation, and Moral Obligation," I discuss the phenomenon of hypermoralized deliberation or rumination and its implications for a suitable ethical theory and decision procedure for the depressed. I argue, following psychological realist critics, that impartial or impersonal ethical theories aim toward an impossible goal of eliminating moral luck and so justifying an absolute measure of ethical merit or credit and blame as the foundation of pure justice. This move tends to rely on a strict distinction between moral obligation, which has a special overriding status, and prudential interests. This strictness may appeal to depressed persons who have lost trust in their own capacity to evaluate their ethical worth and desire an absolute measure of this worth that they can trust. But depressed persons, in deliberation, compulsively efface the distinction between the moral and the prudential: every possible act seems to have an ethical status, such that every failure to act seems evidence of their ethical unworthiness and every act taken seems to fall short of a vague but absolute ideal. I identify the practical effacing of the moral-prudential distinction with an experience of self-alienation, or the loss of a sense of agency in the experience of forming intentions. In losing the sense of "mineness" of one's personal goals and projects as an intuitive epistemic guide to weighing options in intention formation, one becomes more likely to spiral into a moralizing rumination that flattens experiences of value by presenting all personal desires as impersonal and as more or less worthy ethical imperatives.

I thus conclude that an appropriate ethics for the depressed is one that is "a reliable guide" rather than strict or partial rather than impartial. It acknowledges, first, the

38

possibility and likelihood of demoralization and hypermoralized deliberation and, second, the ineliminability or necessity of moral luck at the level of will or character. Its priorities are persistence over self-transcendence, consistency over necessity, the provisional over the certain, and more streams of motivation over better-justified authority. This amounts to a broad acceptance of motives as possibly justifiable reasons and the necessary imperfection of the ethical apportioning of blame, and a broad rejection of a hierarchy of especially moral over prudential considerations. The question becomes what account of reliable or partial ethics is best suited to depressed persons.

The central challenge is that an ethical theory that takes the desires of depressed persons as a starting point must ultimately justify self-transcendence, for two reasons. First, I affirm that an ethical theory is a "recognizable ethics" only if it may possibly justify the rejection of one's initial or prior desires or values. If an ethics for the depressed, in being partial rather than impartial, effectively elevates partiality over impartiality in every case, then it risks merely being an affirmation of prior interests and affinities and failing to be an ethics in any meaningful sense. It would be mere dogmatism, theoretically patting depressed persons on the back just for being themselves. This is not helpful or necessary coming from a philosophical theory. Second, even if an ethics for the depressed asserts the value of personal motives and projects, it also must explain and justify the possibility and value of transcending these personal motives and projects, if only because a depressed person's deepest desire is likely to be the transcendence of existential guilt as the cause of failures to pursue some ethical goods.

In Chapter 5, "A Value Ethics of Engagement," I argue that an ethics for the depressed, though partial or merely reliable rather than impartial and absolute, may

justify a contingent ranking or hierarchy of values, ranked according to their tendency to produce a feeling associated with engagement: a sense of stability. I distinguish engagement from elation or excitement by claiming that it includes a sense of stability or connectedness with another person or a greater purpose, like a sense of belonging or being at home. I argue that depressed persons, in being aware of their vulnerability to deep senses of precarity and isolation and averse to those experiences, are inclined to seek out a sense of stability through engagement. Moreover, I argue that the "less partial" value paradigms of compassion or kindness and fairness are, overall and in the long run, better sources of the sense of stability than the "more partial" value paradigms of, e.g., loyalty, deference to authority, and sanctity. This argument is in part based on a conceptual claim, that the latter imply an in-group/out-group distinction while the former do not, and in part based on an empirical claim that values that do not imply an in-group/out-group distinction are more productive of a sense of stability overall and in the long run than those that do. I thereby justify, for depressed persons attuned to their desire for a sense of stability, the greater "height" of the values of compassion and fairness over the "lower" values of loyalty, deference to authority, and sanctity. This contingent hierarchy of value may ground an imperative of self-transcendence and transcendence of one's in-group in a partial or merely reliable ethics.

This argument stands or falls on the claim that "less partial" values are overall and in the long run more productive of a sense of stability than "more partial" values. I thus take the strongest objection to a value ethics of engagement to be what I call the problem of the walled garden: depressed persons attracted to stability and belonging and averse to precarity and isolation may, in insular communities, prefer values that prioritize their in-

group over out-groups out of fear of ostracization or shaming, which would induce a sense of precarity and isolation. A depressed person as I have defined them might, for example, feel a deep sense of belonging as they join their community in joyfully inflicting terrible cruelty and injustice on an out-group. It may seem that according to a value ethics of engagement, this cruelty and injustice would be not only permissible but laudable in its generation of a sustainable sense of stability, thereby completely contradicting and overturning its ostensible hierarchy of values. This result would then be a total defeat for an ethics for the depressed and a strong argument for an impartial ethics, even if I am right that the latter is doomed to fail to refute moral luck and so to disappoint or devastate the depressed.

Here, I present two separate answers, one for the already autonomous and another for the aspiringly autonomous, which is to say, depressed persons. For the already autonomous: I believe that the problem of the walled garden motivates the necessity of a politics of engagement to contextualize an ethics of engagement before the latter may succeed in being generalizable beyond the depressed. A politics for the depressed or a generalized politics of engagement would explore the relation between 1) value pluralism and political pluralism, 2) the heart as the set of felt values and the person as the bearer of dignity and rights, and 3) the sense of stability and solidarity. I conclude, in Chapter 6, by pointing toward a "proto-politics" in which an act motivated by and justified in terms of integrity may transcend communal values, and so risk alienation from a community, and achieve solidarity. If a partial or reliable ethics may, in justifying integrity as a contingent but consistent good in response to qualitative ambiguity or opacity and in incorporating

41

higher values, thereby justify the transcendence of in-group values and the achievement of new solidarity, then it may overcome the problem of the walled garden.

My answer to depressed persons is different and simpler: I argue that the depressed person already possesses a tendency to transcend walled gardens, not because they are intrinsically attuned to higher values but because they are prone to alienation or a loss of sense of belonging and thus cannot count on dogmatism or conformism to provide sufficient motivation or reasons for going on.

In Chapter 6, "The Call of the Other," I redefine the depressed person's tendency toward altruism as a tendency to experience the call of the other in situations of normative breakdown. The call of the other is psychologically simple: it is a combination of sympathy and alienation, where alienation is a loss of the sense of oneself as an individual and sympathy presents itself as individuating, as a kind of evidence that one's unique existence matters, and so as a solution to alienation. But when it arises in a context of normative breakdown, it becomes phenomenologically complex: as a motive for ethical action that may be more easily and successfully taken to be a good enough reason for ethical action, it may seem to be the essence of ethical responsiveness as such. However, in contrast to phenomenological ethics of alterity, I claim that the call of the other is ultimately a cause or motive of action, not intrinsically a reason for action. It is just comparatively straightforwardly taken to be a reason in situations where most other reasons seem to fail or fall apart. The main benefit of my conclusion for the depressed is that it situates alienation as contingent, not necessary, in revealing one's vulnerability or possibilities of ethical responsiveness. I show how a depressed person may begin with personal and direct altruistic action motivated and justified by appeal to the call of the

other, move on to personal and indirect altruistic action in which the depressed person is anonymous to the recipients, and then move on to impersonal and indirect altruistic action in which the recipients are also anonymous to the depressed person, all in the context of a partial and merely reliable ethics. It is thus unnecessary either to restrict ethical responsiveness to personal and direct responses to others in the moment or to attempt to abstract some transcendental condition of responsiveness from such experiences as a general ground of ethics.

The call of the other is a manifestation of existential guilt and a contingent, not necessary, condition of 1) ethical responsiveness and 2) the coming to awareness of vulnerability or the dependence of autonomy on care and trust. But its possibility is also the possibility of motivating and justifying self-transcendence and communal transcendence. Sympathy alone might easily be restricted to one's group or faction. But in a context of alienation or the loss of a sense of belonging, to experience sympathy as a solution and response to alienation is more likely to mean leaving one's walled garden in search of something more, some better reason to go on.

It follows that existential guilt gives the reason for its own transcendence. The call of the other is the beginning of an awakening to existential hope and the possibility of finding better, more motivating and better justifiable, reasons. Existential guilt is not itself responsive to reasons, but it is both 1) manageable indirectly through effective responses to demoralization and hypermoralized deliberation and 2) motivates inquiry into reasons by the aspiringly autonomous. In framing existential guilt as the basis of an ethical life, one that is eased without being eliminated by a different way of ethical thinking and ethical life, I aim to show that existential guilt is, for all its pointless pain,

basically good. I am to *make* existential guilt good, to retroactively justify its otherwise senseless presence, by showing that it may make depressed persons good. The trick is in going on.

I conclude this introductory section with a short piece that I title "Ethics for the Depressed 101: Arguing with S," which is written at an Intro to Philosophy level and presents the throughline of my dissertation as a response to a depressed friend.

In future work, I hope to add to the contents of this dissertation further reflections on practical or group identity as pre-intentional, on virtue ethics, and on political philosophy. I might also extract my original research into 1) moral psychology and quasi-beliefs, 2) the distinction between existential feelings and existential dispositions, 3) qualitatively ambiguous situations and qualitatively opaque problems, and 4) phenomenology of value. While I take all of these to be contributions to the scholarship in question, I acknowledge that they may distract from, even as they argumentatively support, the main ethical argument.

Ethics for the Depressed 101: Arguing with S

I have a friend, S, who also struggles with depression. They are intelligent, charismatic, funny, and generous. I love them very much. But I think that they have difficulty following through on long-term goals or projects, and I worry that they are dissatisfied with, and feel undeserved shame about, their struggles. Their days may be disrupted and their time consumed by episodes of deep listlessness, during which they describe themselves as feeling unmotivated and yet bored, both unable to act and

dissatisfied with passivity. They have sought treatment for depression but, thus far, it has not been sufficient to forestall these episodes.

Recently, I had an argument with S about how they might respond to these episodes. I claimed, basically, that exploring different treatments might be worth trying despite the possibly considerable cost and effort involved. Their primary objection was not to my weighing of costs and benefits. Rather, they objected to the notion that the problem was at all internal rather than external to them. They are hopeless, they claimed, because the world is without hope.

I will call my rendering of their argument, in S's honor, the "I Am Going to Become the Joker" argument, or The Joker Argument for short. In the hope that it will entertain S, and readers generally, I now present it in the style of a generic but melodramatic comic book villain:

"Look at the world around us. Inequality of opportunity and outcome is extreme. Frauds and grifters prey unobstructed on anyone who relaxes their vigilance. Disease runs rampant. Soon, climate change will ravage the earth until we spiral into war and famine. Politicians do nothing and will continue to do nothing as they feed from the festering trough of the wealthy.

In this world, still, I have tried to pursue my ambitions, to realize my dreams. I have failed. Of course, I have failed. This world is designed to deny our success. I am not to blame. Yet in one last spiteful mark of this world's injustice, I cannot help but bear the shame.

You say that I am compelled by depression to some irrational conclusion. But there is nothing irrational about my response to this world. This world is hopeless, and

45

my hopelessness is a rational recognition of its hopelessness. I am right to be hopeless. You are wrong to hope."

One may reasonably find this argument either moving or silly. But I think that it is powerful in an unsettling way, in part because it is premised on a pessimism about the world that may more accurately predict the future than optimism. It is also fatally flawed by a missing premise. Because S is a logical person, my case to them benefits from formalizing their claims.

Figure 1. The author's rendering of S. (The author has not seen the movie, just its memes.)

i.   S's Argument

A formalized version of their argument might look something like this:

1) **Premise 1**. If, for S, success in their hoped-for goals is impossible, then hope is, for S, irrational.

2) **Premise 2**. In this world, only a select few may succeed in their hoped-for goals. For all others, the "unfortunate," success in their hoped-for goals is impossible.

46

3) **Premise 3**. S is a member of the unfortunate, not the select few.

4) **Conclusion 1** (*from Premise 2 and Premise 3*). For S, success in their hoped-for goals is impossible.

5) **Conclusion 2** (*from Premise 1 and Conclusion 1*). Hope is, for S, irrational.

6) **Conclusion 3** (*from Conclusion 2*). S should not hope.

Let us not yet extend this argument beyond S. Let us also assume, for the sake of argument, that all three premises are true. Note, however, that there are good reasons to doubt them all.

Premise 1 might seem questionable or at least overly simplistic. Its truth depends on what one means by "rational." But if we understand the term "rational" to imply the possibility of success, then Premise 1 will be tautologically true. Let us take the term "rational" to imply the possibility of success, at least in the case of what is called means-end or instrumental reasoning.

Premise 2 suffers from being both excessively absolute and excessively vague. Set aside that "success is impossible" is a sweeping claim that is at least difficult, if not itself impossible, to justify compared to a more moderate claim like "success is highly unlikely." Set aside too its morbidly hierarchical vision of the world, which admits of no alternative egalitarian communities where better equality of opportunity and outcome may exist. If S hopes for, say, wealth in an economically unequal world, then the premise may seem intuitive precisely because of these otherwise questionable absolutist assumptions. But the extent to which S's hopes will be obstructed by the world will depend on what specifically S hopes for. That is, Premise 2 is also vague about if some

47

hoped-for goals might be more successfully pursued than others, even in an unjust or what S calls a "hopeless" world.

This problem of vagueness will then affect the truth of Premise 3, because whether S is personally likely to succeed in S's hoped-for goals may also depend on what those goals are.

But let us grant all three premises. We have a better target. Before we proceed by developing a counterargument to S, however, let us pause to ask: what does S mean by "hope"?

Hope in S's sense seems to be what I would call standard, goal-oriented, or "intentional" hope. It is 1) a motive to pursue goals, 2) can be taken as a reason to pursue goals, and 3) involves emphasizing, either by choice or inclination, the possibility of success over the possibility of failure. A hopeful person, in this sense, is hopeful because they either 1) believe that they will succeed in their hoped-for goals or 2) simply do not dwell overmuch on the possibility of failure to achieve their hoped-for goals. They may even, when deliberating or struggling with self-doubt, consciously reflect on their own hope as a reason for going on.

I note that "intentional," in the sense of intentional hope, has an unconventional meaning borrowed from phenomenology and linguistics: it means "directed towards something" or "about something." For example, fear or intentional hope are fear or hope *of* or *for* something. But we can understand, say, pain as just being a sensation that, even if located in specific parts of the body, is not "about" anything. Pain has causes, like injury, but the experience of pain is not "directed towards" or "about" its causes in the

same way that fear or intentional hope are "directed towards" or "about" their objects. This definition will become relevant soon.

Now that we have a preliminary definition of hope for S, let us return to S's argument. Conclusions 1 and 2 seem to logically follow from the premises. According to Premises 2 and 3, S, as an unfortunate, cannot succeed in their hoped-for goals and so, according to Premise 1, hope is, for S, irrational. But does Conclusion 3 logically follow from Conclusion 2? Does it logically follow from hope being irrational for S that S should not hope?

No, it does not. There is a missing premise:

5.5) **Premise 4**. If hope is, for S, irrational, then S should not hope.

This is not just any old missing premise. It is the crucial move to the normative, to what one *should* do, from the descriptive, or what *is* the case. Premise 4 is logically necessary to make the move from the descriptive ("is") Conclusions 1 and 2 to the normative ("should") Conclusion 3.

To illustrate the normative, consider the normative implications of Premise 1. The concept of "rational" versus "irrational" depends on something like a concept of a "good reason," or a normative notion of "acceptability." Reasons are either "good" and *should* be rationally accepted or "bad" and *should not* be rationally accepted. The application of this norm of acceptability to specific reasons then depends on some criterion of acceptability for specific purposes. For example, a reason might be good if it is accurate in what it describes or predicts.

Now, we have already accepted Premise 1 for the sake of argument. We are not debating the criterion for what amounts to a good reason to take hope to be *rational* or

not. What we are missing is the criterion behind Premise 4. We have not said what counts as a good reason to take hope to be *imperative* or not. How do we know that if hope is irrational, S therefore should not hope? That is, how do we know that Premise 4 is true? Where is the argument for Premise 4?

Premise 4 is the missing claim, but it does not include its own justification. Behind the claim "if hope is, for S, irrational, then S should not hope" must be some criterion about what counts as a good reason for what S, or anyone, should hope. But that criterion is absent here. It is not yet included in our set of premises. It is not part of our set of facts about the world.

Premise 4 is where I would strike. Consider a counterexample: the existentialist as depicted in Albert Camus's *The Myth of Sisyphus*.

ii.   The Existentialist's Argument

The existentialist believes, basically, that the world is hopeless, just as S says: all our acts will be forgotten, all our works will turn to dust, and all meaning is ephemeral, ever slipping through our fingers. Even so, the existentialist chooses to act. It is precisely because there is no external reason to act that the existentialist, in choosing to act, must therefore be choosing entirely for themselves. The experience of choosing to act without an external reason to justify action becomes the experience of absolute freedom, which they take to be valuable. Thus, absurdly, heroically, the existentialist pursues their goals, like the mythical Sisyphus pushing a boulder up a slope forever, repeating endlessly, achieving nothing.

Maybe it would be a mistake to call the existentialist's reason for acting "hope." Arguably, the existentialist is precisely acting in the absence of hope for anything. According to our definition of standard or intentional hope, it involves an emphasis on the possibility of success over the possibility of failure. But the existentialist does not believe in the possibility of success. By that definition, it is incorrect to call them hopeful.

Nevertheless, the existentialist is acting toward goals for a reason. S seems to be framing "hopelessness" as a reason for inaction. Thus, again for the sake of argument, let us define a different kind of hope behind the fact that the existentialist acts in a "hopeless world." Let us say that the reason for the existentialist's actions is some radical, transcendent, existential version of "hope." This "existential hope," in contrast to intentional hope, is not hope for success or any specific achievement. Instead, it represents the existentialist's broad assertion, for them unjustifiable on the grounds of any external reason, that their own freely chosen action is valuable regardless of its success or failure. To try and fail, over and over, forever, may then be understood as guided by and expressive of existential hope.

I said that S seems, when discussing hoped-for goals in the premises of their argument, to be discussing intentional hope. But, thinking back, to say something like "the world is hopeless" is an odd claim in the context of intentional hope. Persons, not worlds, have hopes. Perhaps S just means to say: "the world makes it irrational for me to intentionally hope." Then S is simply restating Premise 1 and eventually making a logical leap via the absent, unjustified Premise 4.

But there is another explanation. S's concerns with intentional hope may be a symptom of their concerns with existential hope. That is, though the premises of their

51

argument concern intentional hope, S's real concern may be one that they are not describing or cannot describe: they are aware that they lack the existentialist's existential hope and that if they are to go on in a world in which success is impossible, something like existential hope may be necessary.[22]

My point is not that the existentialist's argument is right and that S's argument is wrong. My point is that it is *not obvious that S's argument is right and that the existentialist's argument is wrong*. S argues that the world gives them their reasons to be hopeless. But for this argument to hold, the existentialist, whom I have argued is hopeful in a hopeless world, must somehow be refuted. S has not presented this refutation because S's premises only refer to intentional hope.

I have chosen this example strategically. I have not merely targeted Premise 4 because I think it is the weakest premise. I have targeted it because I think S will be sympathetic to the existentialist who is its counterexample, and I am playing on their sentiments. I think that they will admire the existentialist, even desire to be like them, yet feel powerless to do so. That is what I want to direct their attention towards. Why do they feel powerless? If there is no *reason* that the existentialist is wrong to hope against hope, what is the *cause* of S's feeling of powerlessness to, say, emulate them? Both S and the existentialist seem to see themselves as being powerless to make the world a hopeful place. But why, then, does a hopeless world, which seems for the existentialist to be a *reason to act*, seem obviously to S to be a *reason for inaction*?

---

[22] In a future revision of this piece, I intend to formalize the existentialist's argument and attempt to show that even if S's argument involves true claims about intentional hope, these claims are not necessarily true of existential hope, which is the sort of hope that I surmise S truly desires. I am grateful to Claire Pickard Mairead for these ideas.

To reformulate my original claim to S: even if we grant that all the causes of S's feelings of powerlessness are external (the hopelessness of the world) and none are internal (like, say, brain chemistry or trauma), the cause of the feeling of powerlessness *seeming to obviously be a reason for inaction* must be at least partly *internal*. This follows because the existentialist and S differ in their responses to their equally hopeless external worlds. If causes of action are either internal or external and there is no external difference between the worlds of S and the existentialist, and yet their actions differ, it follows that there must be an internal difference between them. It may be a problem that the world is hopeless. But it is also, separately, a problem that S is hopeless.

Now, I must be very clear here. First, S did not choose to be hopeless and is not to blame for their hopelessness. I assume that S cannot simply choose to become more hopeful and call it a day. To claim or even to imply otherwise is cruel and unfair to S.

Second, I am not claiming that the internal cause of S's inaction is "more important" than the external causes. We can agree that the hopelessness of the world, the unjust way it undermines the hopes of the unfortunate, is a serious problem. My point is simply that as an external problem may require external solutions (by which I mean, actions to change the world, if possible), so an internal problem may require internal solutions (including indirect approaches like, say, different medication or therapies). Furthermore, individual responses to internal problems (like seeking different medication or therapies) may be more likely to succeed than individual responses to external problems (like trying to change the world on one's own). This claim, in stating that some hoped-for goals may be more likely to succeed than others, conflicts with the absolutism of S's Premise 2. But I think it may still seem reasonable to S.

Third, as I think their persuasive argument shows, S is very reasonable. Yet they *still* compulsively and unknowingly assert Premise 4. They erroneously take their assertion of an ostensibly factual claim, that it is impossible for them to succeed, to be a reason for inaction, despite the valid counterexample of the existentialist. Let us examine this point more closely.

iii.   Leaps and Compulsions

Now, it is typical to take facts to obviously be reasons, at least in practice. I do so all the time. If I perceive a dangerous object to be hurtling at me at dangerous speeds, I do not have to reflectively conclude that I have a reason to get out of the way. My mind takes a shortcut around conscious thought to protect me. I may be leaping aside before I even understand what I am doing. Moreover, although I did not choose to leap, I will judge afterwards that it was good to leap. This leap is consistent with my desires and beliefs regarding my own self-preservation.

Here is a useful way of understanding practical reasoning. If we understand "acting rationally" in a purely reflective sense, to mean "reflectively endorsing a reason for action," then my instinctive leap would not be a rational act. I did not, in that moment, choose to value my self-preservation and then deliberately act on that choice. But if we understand "acting rationally" to mean, in a more practical sense, "acting in accordance with reasons that one would reflectively endorse," then even an instinctive act like my leap is rational. True, my mind took a shortcut from the factual claim of "perceived threat" to the action of "avoidance" without deliberating about what that *fact* is a *reason* to do. But even if my leap is not "from" reasons, because it was instinctive and not

54

reflective, it is "in accordance" with reasons of self-preservation that I would reflectively endorse (for example, that I would judge afterwards to be good reasons). Thus, though I was *caused* to leap, my leap is fairly called rational.

Note that I have distinguished between causes and reasons. Outside the context of ethics, it is typical to use the terms "cause" or "motive," on the one hand, and "reason," on the other, interchangeably. The interrogative word "why" is ambiguous between causes and reasons. Both a cause and a reason are "why" something might happen. If I am asked "why" I leapt aside, I will explain that I was dodging the object. It will not matter in this context that I was compelled by instinct, a causal force, to move and that I did not choose, by reflectively endorsing a justifying reason, to move. My answer is an appropriate and acceptable response.

But imagine that instead of leaping, I am hitting someone, and when I am asked why, I respond, "Because I was angry." This is an explanation, but it is not an excuse. My listener may understand my meaning, but it does not follow that they agree that my action was acceptable. In this ethical context, which is the *normative* context of reasons that is the primary theme of ethics, the distinction between causes and reasons is essential and necessary. Anger may be an understandable *cause* of or *motive* for someone hitting someone else, but it is not therefore a *reason* to accept them doing so, let alone a good reason.

Now imagine that instead of leaping to dodge, I am leaping seemingly at random. What if this leap is not consistent with my desires and beliefs? That is, what if I do not want to leap and do not think it is a good idea, but find myself doing so anyway? I will then find myself having already acted in ways contrary to those desires and beliefs. I will

have to react not just to the world but to myself, attempting to correct my own actions after the fact so that they conform to my desires and beliefs. I will have to move back to my previous location after cleaning up whatever mess I have made by leaping in some random direction at some random time. This will require additional effort from me and will likely involve some embarrassment, even if (I dread to even mention the possibility) no one has been hurt.

What if I find myself leaping, inconsistent with my desires and beliefs, over and over? I will then constantly be reacting to my own actions, struggling to resist myself like my own adversary. Eventually, even as I consistently perceive the threats to me as being from without, I may come to suspect that the threats to me are actually coming from within, that some alien force has already infiltrated me. My presumptive boundary between an external world of possible threats and obstructions and an internal world of freedom and security will seem to degrade. I will become at once hypervigilant and exhausted. No amount of vigilance will be enough to forestall the activation of a cognitive shortcut that takes over before I can think.

In the context of depression, I will call this sort of cognitive shortcut a "compulsion." In other contexts, it might be called an "impulsion." The impulsive move from the fact of a threat to a reason to evade seems useful for rapid response in self-preservation. But imagine that a person who is otherwise reasonable may experience the presence of a threat at any time, regardless of if one is present. The self-preserving reaction of aversion will trigger, again and again, and every time this person will have to ask: is something out there? This time, what if my instincts are right? I'd better wait. I'd better hide. I'd better just lie still for a while. After all, what is the alternative but inching

56

forward, my every synapse crackling and every nerve straining, telling me over and over I'm in danger from a threat that must be there even if there is no other evidence?

In this story, the mind makes just one error: either some assertion or some sense that a threat is present, one that makes the possible failure of prospective action seem both punishing and practically guaranteed. This threat may come to mind in any number of ways. It may stick to any object of one's perception or reflection. It may be fear of the shadows, or fear of getting out of bed, or fear of the shame of failure, or fear of being noticed and called a coward, or fear of what one's fearfulness says about one's own character. It may seem to switch between these fears or to be all of them at once. This is because it does not matter what the fear is of. To turn to metaphor: the mind has already decided that there is a threat. The heart is left to sort out the rest.

Maybe it would be a mistake to take S as an exemplar for depression. I have, admittedly in jest, presented S as villainous. But depressed persons are not generally villains. If anything, most incline toward self-sacrificial altruism. I take my caricature of S to represent one possible, and especially defiant, response to a general sense of precarity and isolation, one that may also manifest as anxious paralysis or retreat. If depression may be fairly, if very loosely, defined in reference to this deep sense of precarity and isolation, then my argument with S is useful in identifying compulsions most significant in depression. A depressed person may not identify with S. But they may empathize with S or have had such thoughts themselves at least once.

Let us identify the compulsions most significant in depression, then. First, of course, there is Premise 4, the jump from the impossibility of success to the imperative of inaction. But there is also Premise 2, the initial assertion of the impossibility of success.

As I noted, "impossible" is a sweeping claim. I venture that S does not really think that successfully acting toward their hopes is impossible, or at least, they do not *just* think so. I suspect that they already have thoughtful plans for action that involve reasonable estimations of success. But even if they do have these plans, Premise 2 overrides them, effectively rendering them irrelevant. Thus, S may come to both believe and disbelieve in the possibility of success.

The absolutism and vagueness of Premise 2 is, like Premise 4, the product of compulsion. It is like the threat that seems to come from everywhere and nowhere, at once absolute and vague, disrupting action and destroying hope.

iv. Ethics for the Depressed

The intuitive leap from facts to reasons may be self-preserving, as when I dodge. Thus, it is not necessarily destructive. It may even be the beginning of ethical reasoning as such. But S is making two specific leaps, first from the unlikeliness of success to the impossibility of success, and then from the impossibility of success to the imperative not to try. The point is not that S is wrong. The point is that S, who is otherwise reasoning carefully, is making these leaps anyway. Thus, S's internal problem has two components: it is the compulsion to leap first from the seeming unlikeliness of success to the seemingly certain impossibility of success, and then from the seemingly certain impossibility of success to the imperative of inaction.

But now we have a new problem. Let us assume that all that I have said is correct. S is thus under a compulsion to assert Premises 2 and 4. So what have I achieved by arguing with them? Haven't I suggested that S already believes that Premise 2 is false?

Yet S has asserted Premise 2 anyway. If S's self-refutation is ineffective, why would my refutation be effective? Thus, my error is not internal to my argument. My error is in attempting to argue with S. Again, S is a very reasonable person. But Premises 2 and 4 are not responsive to reasons, theirs or mine. In arguing, I am being stubborn at best and cruel at worst, myself compulsively philosophizing.

Point taken. Yet I think it is possible for this argument to achieve at least two things:

1) **Exhortation through Description**. S concluded that the world gives them sufficient reason for inaction. I have demonstrated this to be false by refuting Premise 4. Even if S compulsively asserts Premise 2 and 4, if S recognizes them *as* compulsive, their attitude toward these premises may change. S may come to view them as internal threats to hope and thus acknowledge that there *are* internal, and not just external, threats to hope. Then, perhaps S will be motivated to seek new methods of easing their compulsive power. Even if S cannot choose to be more hopeful, there may yet be treatments or mental techniques that will ease their episodes of hopelessness, thereby enabling them to better act in accord with their other desires and beliefs. Now, the idea that I may "motivate" S might seem strange, since one might think that S lacks any desire to seek solutions to their internal problems. But in my telling, S's problem arises from two specific compulsive mental leaps. Desire was not mentioned. It is thus entirely possible that S retains all the desires that provide them the motivation to pursue and value their goals. S's problem would then only seem to be a "lack of motivation." Rather, S *would be motivated but*

effectively unable to act due to the compulsive and absolute assertion of the impossibility of any action's success. This would explain why S, in depression, is bored and dissatisfied, yet does not act.

2) **Exhortation through Prescription**. This is a riskier bet. Perhaps S will only feel frustrated with me or with themselves as our argument collides impotently with Premises 2 and 4. Perhaps S will feel disrespected, victimized, betrayed, or humiliated by the experience that I will seem to have inflicted. But there is another option. S may also feel respected because I am taking S seriously. I am treating S not merely as sick or confused but as making an argument worthy of consideration. Arguing with S will likely not be enough to change S's mind, because some of S's premises are the products of compulsion. Worse still, it is possible that by showing S that there is a hypothetical way forward from their position (that being the indirect rejection, through treatment and mental techniques, of Premises 2 and 4), they will only feel more isolated. They may feel that although this hypothetical way forward exists for others, they themselves, somehow, cannot proceed with it. But I am placing my hopes on another possibility: that through reasoning, S experiences themselves as an essentially reasonable person. S is *mostly* responsive to reasons. More importantly, S *wants* to respond to reasons. Otherwise, S would not bother to argue fairly with me.

If someone wants to respond to reasons and is already at least somewhat responsive to reasons, then whatever they might think, there is hope for them. They may still find

reasons to go on. They need not wait for the world to give them those reasons. In truth, they already have them.

S is subject to two specific compulsions: 1) to perceive an absolute and vague threat that causes disruption to action and 2) to take this cause as also being a reason for inaction. The first half of my dissertation deals largely with the first point: it is a metaethics for the depressed, or "ethics as a reliable guide" as a response to "demoralization" and "hypermoralized deliberation." There I challenge what I call the Stocker-Smith account of depressive loss of motivation as being a loss of desires and argue instead that it involves the defeating presence of what the phenomenologist Matthew Ratcliffe calls "pre-intentional" mental states, a category that I redefine and expand to include second-order "quasi-beliefs" and habits of feeling, that interfere with intention formation and action despite the persistence of desire.

The second half of my dissertation deals largely with the second point: it is a normative ethics for the depressed, or a "value ethics of engagement" that is a "gentle perfectionism" of "contingent value ranking." By demonstrating in the first half that depressed persons retain their desires, I become able to premise a value ethics upon what I call the consistent desire for a "sense of stability" in response to experiences of precarity and isolation. My earlier metaphorical distinction between the "mind" and the "heart" (which is not literally a claim that the heart is not part of the mind) develops into a concept of the heart as the set of "felt values" or intuitive value paradigms that are themselves pre-intentional states or dispositions. I have thus attempted to structure a complete ethical theory, integrating plural philosophical traditions and founded on the phenomenological category of pre-intentional mental states, in response to the presence

of two specific compulsions experienced by an otherwise reasonable interlocutor. I put an orthodox style of philosophy in service of an unorthodox agent: one who is "aspiringly autonomous."

I have done this knowing that, in the context of depression, theory and argument only go so far. My descriptions of the causes of depressive inaction are not themselves reasons for depressed persons to act. My prescriptive case for endorsing ethical reasons for action that might more successfully align with depressive motives for action remains vulnerable to the same depressive sense of isolation that it aims to ease. A depressed person may always respond to me, "What you say may be true for others, but I am certain, somehow, that it is not true for me." That response is resistant to refutation. Their certainty might not yield to any argument I muster.

There is one last problem. I fear that S's pessimism is accurate. Even if S becomes able to hope, that hope will likely be frail. If the world is too cruel or unfair, their trust in it may be betrayed and their hope may be destroyed. Changing S, or many Ss, may not change the world.

Still, I am hopeful. I am hopeful that my theory may give hope, that hope may give way to trust, and that, through trust, the world may change, becoming less hopeless and more trustworthy for us all, S included.

CHAPTER II: DEPRESSION AND MOTIVATION

Depression is characterized by a loss of motivation, but how to understand this loss in philosophical terms remains obscure. A depressed person may experience "loss of interest or pleasure in nearly all activities" (DSM-5, 163), suggesting an absence of motivating desires. In Michael Stocker's "Desiring the Bad" (1979) and Michael Smith's *The Moral Problem* (1994), which cites Stocker's essay, this loss of interest is the whole story about depression's effect on motivation. In Stocker's words, "maladies of the spirit" are "extinguishing all desire for good for oneself" (1979, 745); in Smith's uptake of Stocker, depression may "remove any desire" to act on rational beliefs (1994, 135). Loss of desire means loss of motivation, period. This story about depressive loss of motivation does not assign any clear relevance to another common symptom of depression: "feelings of worthlessness or guilt" (DSM-5, 163). As the language of "guilt" implies, these feelings tend to present themselves as having a normative dimension. A depressed person typically feels that they should not be how they are and that this is somehow their fault.

But these two symptoms, loss of interest and feelings of worthlessness or guilt, can also appear to the one experiencing them to be related. Matthew Ratcliffe (2015) quotes an interviewee who "states that depression itself is 'a sign that I'm not what I should be'" (135). Even if a depressed person's goals have no clear ethical status (say, writing some philosophy), their own loss of motivation to pursue those goals may seem to affirm their own lack of ethical worth. This intuition of worthlessness is mistaken. But the intuition that there is some relation between loss of interest and feelings of worthlessness may be correct, even if the feelings of worthlessness themselves are untrustworthy. If a depressed person's loss of interest and feelings of worthlessness or

63

guilt are connected, then they might jointly affect motivation, complicating the story about depression and desire just told.

I share the intuition that loss of interest and feelings of worthlessness or guilt are connected: that connection appears to me as ambiguous yet pervasive, an incoherent tangle. I sense that because I have lost interest in my tasks, I am failing to meet my obligations and so am worthless or guilty; yet I also sense that it is because I am worthless or guilty that I feel that my situation is hopeless and am discouraged from acting. If a depressed person's feelings amount to or result in something like a proposition (like "I am worthless or guilty"), then the vocabulary of desires and beliefs, common to moral psychology like Smith's, can clarify what is incoherent in these intuitions. The first is a Humean formulation that suggests a loss of desire (interest in tasks) leads to a belief or belief-like proposition (that one is worthless or guilty). The second is a non-Humean formulation that instead suggests that same belief or belief-like proposition leads to a loss of desire (being discouraged from acting). If a connection between the two symptoms exists, both formulations of it seem plausible. It is also plausible that both are true at different times, that the "downward spiral" of depression involves some sort of oscillation between loss of desire and change in belief, one leading to the other leading back again.[23] And there is not necessarily one way that depressive

---

[23] In the introduction to *The Upward Spiral* (2015), neuroscientist Alex Korb defines a downward spiral in terms of a relatable example: "We all know what it means to be stuck in a downward spiral. Maybe one Friday night you're invited to a party, but you have a brief thought like *I don't think it'll be that fun*, so you don't go. Instead, you stay up too late on the couch watching television. The next day you sleep in and don't have much energy. No one calls you, so you feel even more isolated, and now you're even less likely to be social. Nothing seems particularly interesting, so you just lie around all weekend. Pretty soon you're unhappy and alone, and you don't know what you can do about it, because every decision feels wrong. This is the edge of what it means to be depressed" (2-3). In the context of this essay, two points stand out from this quote. First: it's not obvious if the "brief thought" described here should be understood as a belief justifying inaction, as it seems at first glance, or if the thought is the product or manifestation of a loss of interest or desire. Second: if this is the "edge" of depression, then it may be helpful to better understand the

loss of motivation begins. But taken as is, these formulations offer incompatible accounts of what, desire or belief, changes first, and thus of how best to understand what is happening when a depressed person loses motivation. This puzzle hints that depression may have more to teach us about motivation than the importance of desire.

How to resolve this tension – which mental state is the depressed chicken and which is its sad egg – remains unclear. It is difficult to parse from experience or testimony how a depressed person's loss of motivation should be foremost understood: as a loss of desire, as a change in beliefs or some other belief-like mental states, as somehow both, or otherwise altogether.

But in moral psychology, where it is standard to understand motivation in terms of belief and desire, the case of a depressed person's loss of motivation involves an additional complication: the fact that depressed persons tend to remain altruistically motivated but become less motivated with regards to their own self-interest. Surveying the available empirical research, Michael Cholbi writes that:

> there is neither convincing behavioral nor attitudinal evidence of indifference to the altruistic moral standards that comprise the core of conventional morality, and in fact, some evidence suggests that depressed individuals are *more* committed to and concerned for moral standards … at the same time that they continue to be concerned with fulfilling their obligations to others, the depressed are frequently indifferent to their own well-being. (2011, 39)

The depressed person's listlessness is generally limited to a specific category of desires, what Cholbi defines as the prudential or "self-regarding concerns" in contrast with conventional morality or altruism (40). Cholbi identifies altruism with "morality" to argue against accusations that depressed persons are "morally listless": "if depressed

---

beginning of this downward spiral, and thus to clarify the first point. Backing away from the edge of a cliff, if possible, tends to be safer and easier than falling off it and climbing back up.

individuals are morally listless," he hypothesizes, "this is likely to be manifest in deficits in altruistic motivation and behavior" (36). But as Cholbi acknowledges, moral philosophers such as Kant and Williams disagree on whether self-regarding moral obligations, like a duty to respect oneself (which a depressed person could plausibly be violating), exist or are notionally coherent. Thus, for the purposes of this discussion, I will provisionally set aside his language of prudence and morality and rely instead on his distinction between self-regarding and other-regarding concerns, with altruism taken as paradigmatic of the latter. In these terms, even as depression may inhibit self-regarding concerns, it may intensify other-regarding concerns, leading depressed persons to "hold themselves to unusually strong standards for altruistic behavior" and "experience intense feelings of guilt when they do not meet these standards" (37-38).

Not only does Cholbi's observation ring true to me but, in what I take to be a second-order symptom of depression, I habitually base a motivational strategy on it: I frequently try to convince myself that my objectives and obligations are somehow altruistic so that they will move me. Even at my most depressed, I can respond to the ethical call of another's suffering. But to my own disgust, I end up moralizing about tasks like getting out of bed or writing philosophy, exhorting myself to act by thinking about how, though I do not feel an ethical call from their prospects, they might eventually serve altruistic ends. This never works. In my frustration, I only feel more guilty and incoherent: my argument is wrong, and it follows that I am self-important and morally confused; it is also right, and it follows that my moral worth is lowered by my failure to act. I now think that, despite the confusion it induces in me, this ineffectual habit of moralizing about my self-interest is based on a naïve but earnest logic: if altruism moves

me and I want to go on, but my objectives and obligations no longer move me themselves, I must make them appear altruistic to me somehow. As the author of this ploy, I never fall for it, but it is kind of clever. It may even be true that there is something altruistic in getting out of bed or writing philosophy. But that truth would make no difference to my lack of motivation: it is only direct and immediate altruistic opportunities that leap out at me and briefly pierce that gloomy veil.

Even if depression is understood as giving rise to mistaken intuitions about one's ethical status or the futility of action, Cholbi's observations suggest that depression has an internal logic that explains how other-regarding concerns could remain intact or be intensified while self-regarding concerns weaken or fade. This story could go: (1) the depressed person feels guilty; (2) this guilt evokes or is otherwise tied to feelings of worthlessness; (3) altruistic concerns present themselves to the depressed person as live demands, perhaps because of their desire to alleviate their guilt and feelings of worthlessness through moral action; (4) yet though altruism thus retains or strengthens its appeal, their guilt and feelings of worthlessness may continue to inhibit their ability to act, especially in their own interest, by, for example, leading the depressed person to predict that they will inevitably fail to effectively act or alleviate their guilt. In this way, a line can plausibly be drawn from guilt and feelings of worthlessness to a loss of motivations to care for oneself and the retention or intensification of motivations to care for others, presenting an alternative to the line drawn from loss of interest to loss of motivation as in Stocker and Smith.

But the limited scope of depressive listlessness, applying as it does to the self-regarding but not the other-regarding, adds another wrinkle to the question of how to

understand depressive loss of motivation. Again, a depressed person is generally listless in matters of self-regard, not regard for others. And again, this restricted scope to listlessness makes sense if the person in question feels themselves, but not others, to be worthless. So, returning to the apparent link between loss of interest and feelings of worthlessness or guilt, the effect of depression on desires (listlessness) might be subject to some sort of framing or filtering by the effect of depression on beliefs (worthlessness) that retains other-regarding motivations but not self-regarding motivations. But the inverse is also plausible. If the depressed person's feelings of worthlessness amount to a belief-like proposition ("I am worthless") that discourages action out of, say, a sense of futility, the effect of depression on beliefs (worthlessness) might instead be framing or filtering the effect of depression on desires (listlessness). The knotty question of how to understand depressive loss of motivation – as loss of desire, change in some belief-like state, somehow both, or something else – becomes further entangled with the question of how best to understand the differing effect of depression on self-regarding and other-regarding motivations within the constraints of belief-desire psychology: say, by regarding some motives as moral and others as prudential, as Cholbi does.

My argument is twofold. First, I claim that depressive loss of motivation, or "demoralization," is better understood in terms of a change in belief rather than a loss of desire: specifically, a depressed person adopts what I call a "second-order quasi-belief," with a content like "success in X is impossible" or "failure to Y is inevitable and catastrophic," that inhibits the development of desire into action. Drawing on Robert Noggle's 2016 research on OCD, I adopt the term "quasi-belief" to represent a belief-like mental state that seems to have propositional content but is compulsive, resistant to

contrary evidence, and persistently coexists with reflectively endorsed beliefs that contradict it. Unlike the first-order quasi-beliefs that Noggle discusses, I argue that a depressed or anxious person acts in line with a second-order quasi-belief, one directed at one's own intentional states: whenever a desire or goal-relevant belief emerges, it is cut short by a quasi-belief that casts the content of that state as futile or threatening. Though these states may also be framed as first-order self-referential quasi-beliefs like "success for me is impossible," the logical connection between such a proposition and loss of motivation is the second-order quasi-belief following from this self-referential claim: if success is impossible for me, then for any action X that I take, success will be impossible.[24] This model resolves the issues that critics raise with the Stocker-Smith account, most significantly that if a depressed person loses desire, it is difficult to understand why a depressed person suffers over their own loss of motivation. Not only must a depressed person contend with the content of their quasi-belief, because that quasi-belief is recalcitrant, any contradiction between their beliefs and quasi-belief is irresolvable through argument, compounding their sense of powerlessness or lack of agency. With this model, Cholbi's claim that depressed persons feel the pull of other-regarding concerns while failing to be moved by self-regarding ones (or being moved by self-disregarding ones) could be understood not in terms of depressed persons being moved by the "moral" and not by the "prudential" but instead as the presence of a quasi-belief like "I am a bad person and must redeem myself above all else," such that altruistic motives translate into action even in the face of a perceived high likelihood of failure or

---

negative consequences. I will define this form of altruism as "personal and direct" and discuss other variants in Chapter 6.

This chapter will show how the quasi-belief model accounts for depressive loss of motivation, or demoralization, while remaining consistent with the model of belief-desire psychology dominant in moral psychology today. However, I will go on to show that there is an equally plausible model of demoralization, what I will call the habit model, for which demoralization is not reducible to intentional states. In this and the next chapter, I will argue that even though this quasi-belief model works as an explanation for demoralization, its limits challenge the preoccupation of moral psychologists with intentional states as conditions of motivation, and I recommend a new emphasis on a category of mental states that, following Ratcliffe, I call "the pre-intentional." The notion of a second-order quasi-belief is highly abstract, but because experiences of depression are so difficult to parse, a model that explains and predicts depressive thinking, feeling, and action may be the best that philosophical psychology can do. The limited capacity to introspect into pre-intentional mental states is part of what lumps them together. Because second-order quasi-beliefs probably cannot be introspectively observed and so must be inferred, one might contest the model with a less fine-grained but behaviorally indistinct alternative, like a habit of feeling very threatened or defeated whenever a desire or goal-relevant belief comes to mind. How best to understand the habit model of demoralization and define the category of the pre-intentional, which will encompass both the quasi-belief model and the habit model, will be the subject of the next chapter. Here, I conclude that the notion I develop of "second-order quasi-beliefs" is not mistaken as an explanatory model, per se, but it is a symptom of an intellectual framework that presumes in advance

70

that a condition of motivation must be an intentional state when that conclusion cannot be justified. What depression has to offer moral psychology, then, is not any lesson about judgment, as in the Stocker-Smith model, but an inducement to study the role of the pre-intentional in moral motivation.

I have chosen the term "demoralization" for depressive loss of motivation to evoke what I take to be the broader ethical ramifications of introducing the category of the pre-intentional to moral psychological accounts of motivation. To be demoralized in this context is to feel that one is essentially unworthy and to feel powerless to do anything about it, that is, to compulsively apprehend one's own ethical status as lowered or inadequate such that no ethical act can seem to restore it. In virtue ethics, the term demoralization has been used to describe an actual loss of virtue or character arising from a failure of moral education or societal support.[25] I am instead describing self-perceived lowered ethical status from the perspective of a depressed person reflecting on themselves, not how the depressed person is perceived by others. Thus, to say that a depressed person is demoralized does not entail that others regard them as immoral or blameworthy. These different accounts, echoing the conventional definition of

---

[25] Annette Baier, in "Demoralization, Trust, and the Virtues" (2009), argues that "the moral virtues regulate" our attitude toward our "mutual vulnerability" and "that the point of such attempted regulation is improvement and maintenance of a climate of trust" (177). She understands demoralization in terms of phenomena that undo our moral education and cultivation by damaging the climate of trust (e.g., trauma, extreme poverty) and so effectively reduce our moral agency or capacity to be moral, and she argues that demoralized persons are less blameworthy for their actions. "Demoralization is a disease of the morally fortunate" (179), she writes, and a "climate of trust must first exist before we can expect the virtues that sustain it" (180): only those who have been morally cultivated and had the good luck to live in a trustworthy world can have that trust and their own characters undone. But Baier divides moral and prudential virtues, which narrows the scope of what counts as demoralization; she seems to regard all anti-trust behavior as vicious (raising the question of the ethical status of, e.g., a climate of trust amongst a higher class premised on the subjugation of a lower class); and she does not distinguish between actual and perceived loss of moral agency, risking moral paternalism toward the poor and suffering. Though I thus do not rely on her account of demoralization, I will rely on and further discuss her ethical accounts of care and trust in Chapter 6.

demoralization as discouragement, both involve a loss of agency: the individual's actual and perceived power to affect the phenomenon, and thus presumably their responsibility for it, is limited at best. But my twist against the virtue ethical sense of the term, away from the actual and toward the perceived, is deliberate.

As Cholbi says, depressed persons do not seem to generally be at risk of becoming less ethical. But there are two issues with this notion, one for ethical theory and one for depressed persons. First, the dilemma a depressed person poses for an ethical theory is mainly not how to make or keep them ethical, given their presumptively intrinsic altruistic inclinations, but rather how much they suffer from the guilt associated with their perceived ethical inadequacy, and thus if it is possible at the level of theory (in contrast to therapy) to develop a way of thinking about ethics that is less likely to cause them pain. A fuller account of how existential guilt affects motivation, utilizing the concept of pre-intentional states, will help to understand what this pain is and how it might be relieved.

Second, there may also be a kind of ethical falling-short distinct to depression: the limiting of ethical possibility to personal and direct altruism, like face-to-face charity, as opposed to either less personal and less direct altruism like organizing a charity organization, or ethical goods more associated with fairness than compassion like promise-keeping. Andrew Solomon (1998) describes calling a friend, hanging up immediately, going to his rooftop, considering suicide, going back to his apartment, calling his friend back, being unable to explain, and her refusing to speak to him for two years. His treatment of his friend is the result of depression. It is also, at least taken out of context, ethically wrong: it is unfair and disloyal, even cruel. It is too simple, then, to say

72

that depression is correlated with altruism. It also may be correlated with forms of unfairness or betrayal like promise-breaking or causing unnecessary fear or concern. These may be neither desired nor intended by a depressed person, as in Solomon's case. But if an ethics for the depressed is to begin with depressive intuitions, which is to say existential guilt, it may have to eventually justify the overcoming or transcendence of unvirtuous inclinations.

In laying the groundwork for an account of a depressed person's pain in demoralization, an understanding of motivation that makes sense of depressive demoralization will sharpen a critique of procedural or formalist ethics as alienating to the depressed and bolster support for an alternative ethics of engagement in Chapters 5 and 6.

1. Depression and Motivation in Belief-Desire Psychology

What is striking about the discussion of depression in the last three decades of moral psychology is how little of it there is. In moral psychology, the question of how to understand motivation is generally linked to the question of how to understand judgment. The central issue concerns if motivation is internal or external to moral judgment, that is, if making a judgment like "doing X is good" is a special case of judgment that, unlike judgments like "Y is red," is necessarily linked to action and implies that one is motivated to do X. I think that contemporary philosophers have tended to concern themselves primarily with judgment and only secondarily with motivation (insofar as it sheds light on what judgment is). It is harder to tell empirically if someone has made a judgment than it is to tell if someone is motivated to do something. Motivation,

conventionally understood, brings with it associated feelings, physiological traces, and resulting actions that can be observed and investigated by non-philosophers: it is plausible to say that we know what it looks like. But even if judgments or acts of predication occur in the brain as physical events, the definitional or conceptual question involved in any empirical investigation – "What are we even looking for?" – is much more salient when the topic is so abstract.[26] Depression has come up in philosophical discussions of motivation when it serves an argument in favor of some account of judgment. But articles challenging those conceptions of depression on their own terms, separate from debates about judgment, seem not to have been taken up or responded to by other philosophers. As we will see, insofar as depressive loss of motivation can be properly understood in terms of belief-desire psychology, it has little to contribute to the debates about judgment with which contemporary philosophers tend to be concerned. In what follows, I will not take a position on judgment or its relation to motivation. Instead, I take up the conceptual question of how best to understand motivation and determine, through a renewed look at the example of depressive loss of motivation, that the framework of belief-desire psychology must be expanded to include the category of the pre-intentional.

---

[26] Paul Redding, reflecting in 2020 on the reception of Hegel by contemporary philosophers, refers to (and tentatively affirms) a common generalization about "analytic" and "continental" philosophy: that analytic philosophers are "more likely to take modern natural science, given its undoubted epistemic success, as providing some sort of model for philosophy," and thus to see the history of philosophy as external to the practice of philosophy, as opposed to continental philosophers who treat philosophy "as *itself* unfolding in history and thus subject to broader historical and cultural forces." I would add to this that if philosophy is seen as modeled off natural science, then philosophy must endeavor to do something that natural science does not already do better. This might also explain why philosophers working in metaethics or moral psychology would prefer to be as conceptual as possible, so to speak: it would be our particular contribution to a broader conversation that includes psychologists and cognitive scientists.

The best-known account of depressive loss of motivation in moral psychology, what I have called the Stocker-Smith account, serves for Smith to support one argument for the distinctness of beliefs and desires or, in his terminology, normative and motivating reasons. The eponymous moral problem of Smith's *The Moral Problem* (1994) centers on the tension between the supposed objectivity and practicality of morality: that is, that morality is supposed to be something we can get right (objectivity) and that to make a moral judgment means to adopt an opinion or attitude with a motivational quality (practicality). For Smith, the objectivity of morality implies that moral judgments are beliefs that are made true by some state of affairs like a moral fact, with no intrinsic motivational character, but the practicality of morality implies that moral judgments are desires with intrinsic motivational character. Smith seeks to reconcile commitments to objectivity and practicality through what he calls a "Humean theory of motivation" and an "anti-Humean theory of normative reasons": he distinguishes normative reasons, which involve beliefs about moral facts, from motivating reasons, which are desires (his "Humean" view) that may be criticized on the grounds of normative reasons (his "anti-Humean" view). The distinction between normative and motivational reasons diffuses the seeming contradiction regarding the motivational character of moral judgment: one may make a moral judgment in the sense of acknowledging a normative reason (having a belief about moral facts) without making a moral judgment in the sense of having a motivating reason (having a desire to act on that belief).

In the context of this argument, depressed persons offer an example of agents who acknowledge normative reasons without possessing motivational reasons:

various 'depressions' can sap desire altogether. The depressive may thus know full well that the rational thing for her to do is, for example, to get up and go on with her life: to go to work, to visit a friend, to read a book, to cook a meal, or whatever. But the effect of her depression may be precisely to remove any desire at all that she has to do any of these things. Having no desire at all to get up and get on with her life, she may therefore simply do nothing – or, at any rate, nothing intentionally. (135)

Smith defines normative reasons as "facts about what we would desire under conditions of full rationality" (161) and understands depressed persons not necessarily as irrational but as subject to a set of desires that are "irrational to the extent that they are *wholly and solely* the product of psychological compulsions." The depressed person of his example "may imagine away the effects of her depression on her desires in forming her beliefs about what, under conditions of full rationality, she would desire that she does when suffering from a crippling depression" and thus "even though she has no desire to get up and get on with her life, she may well believe that she has a normative reason to do so" (155). As a sympathetic case of compulsive irrationality, the example of the depressed person strengthens Smith's argument for his theory of motivation by showing how genuine beliefs about what one should do (normative reasons) might coexist with a total lack of desire to act on those beliefs (motivating reasons).

Smith's distinction between normative and motivating reasons, or beliefs as non-motivating and desires as motivating, is also supported by an unrelated argument based on their distinct directions of fit and "counterfactual dependence." As Valerie Tiberius summarizes in her comprehensive introduction to moral psychology, to say that "beliefs and desires have different directions of fit" is to say that "beliefs are mental states that aim to fit the world, while desires are mental states that aim to get the world to fit them" (2015, 48). Smith frames the same notion negatively, in terms of what constitutes a

counterfactual to said states: "a belief that p tends to go out of existence in the presence

of a perception with the content that not p, whereas a desire that p tends to endure,

disposing the subject in that state to bring it about that p" (1994, 115). What is relevant to

our discussion is not the merits of this claim but that it stands apart from the argument

from depression and that it has been significantly more influential. While Tiberius's text

contains five chapters dedicated to motivation and an extensive discussion of Smith's

"Humean Theory of Motivation," depression is mentioned only once in a later chapter on

well-being (2015, 171). I take this to be not an oversight on the author's part but rather a

sign of how irrelevant depression becomes to philosophical debates about motivation if it

is understood just as a tragic and all-too-pervasive example of the loss of desires, leaving

beliefs relevant to moral motivation intact. When the question concerns the relation

between motivation and the loss of desires, it is not obvious why it would matter how

desires are lost.

Yet criticisms of the Stocker-Smith account of depressive loss of desire made by

authors like John Russell Roberts (2001), Steven Swartzer (2015; 2018), and Iain Law

(2009) suggest an alternative theory of depression's effect on motivation that

nevertheless remains compatible with mainstream belief-desire psychology: depressive

loss of motivation emerges not through a loss of desire to act but through a "quasi-belief"

that successful action will be impossible. I will proceed to develop this "depression as

compulsive quasi-belief" theory through first overviewing Tiberius's account of moral

motivation and then turning to the above authors' distinct theories regarding depression

and motivation. Though I will ultimately challenge the "depression as compulsive quasi-

belief" theory as unjustifiably presuming that conditions of motivation must be

intentional states, I will argue that it is the best account that remains consistent with belief-desire psychology. In formulating and then challenging it, I attempt to bolster my claim that the case of depression shows how belief-desire psychology is an inadequate theoretical framework for understanding motivation so long as it neglects the pre-intentional.

Tiberius distinguishes between a narrow and a broad understanding of moral motivation. Narrowly, philosophers working in moral psychology tend to be concerned with the metaethical question of if or how moral judgments or acts of predication (e.g., "X is good") motivate us. The question of if "moral reasons necessarily motivate people," for example, is a "conceptual question": it primarily concerns how best to understand the concept of moral reasons (2015, 7). But broadly, the question "what motivates us to be moral" includes more than just conceptual investigation (29). As an example, Tiberius cites the enduring question of "the role of emotions in moral motivation" as an empirical question that requires normative assumptions: it is possible to scientifically test if emotions help or hinder moral action if a normative criterion for what counts as moral action has been decided upon. Tiberius advocates for a big-tent approach to moral psychology that includes "normative, conceptual (or theoretical) and scientific (or narrowly empirical)" questions and methods. Still, most philosophical debates on moral motivation, and the views on moral psychology at stake in those debates, understand moral motivation in terms of moral reasons.

Some of this ambiguity arises from the term "motive" itself. As Iakovos Vasiliou points out in his introduction to *Moral Motivation: A History*, "motive" may be understood in two senses: as "teleological," or regarding one's purpose in acting, and as

78

"efficient-causal," regarding what has brought one's action about (2016, 8). This distinction helps, for example, when discussing the relation of emotions to moral motivation: if I become angry when I perceive my friend to be threatened and I act in their defense, that anger might energize me to act, and so motivate me in the causal sense, while my teleological motive would be to defend my friend, a purpose distinct from anger. The term "motive" can thus be clarified to some extent by distinguishing between "reasons for" and "causes of" action. This distinction lends itself to a methodological division of labor, with philosophers being concerned with normative or conceptual questions about "reasons" and scientists being concerned with accounts of "causes." Even so, the discipline of moral psychology involves the entanglement of normative, conceptual, and empirical questions and methods, as in Tiberius's example of the role of emotions in moral motivation. But this teleological-causal distinction helps to make sense of why the contemporary philosophical discussion of moral motivation has been preoccupied with questions pertaining to moral reasons.

As Tiberius explains, at stake in the discussion of moral reasons is the link between reasons and motivation, typically framed as a contest between common interpretations of Kant and Hume over the scope of normative requirements, or more specifically, the possibility of categorical or non-contingent motivating reasons. A Kantian theory of moral motivation espouses an intuitive claim: it is possible for someone to have a reason to do something that they do not want to do, such as a moral duty they find burdensome. But a Humean theory of moral motivation, like Smith's, puts forward the intuitive claim that desires are intrinsically motivational and beliefs are not: first we want something, then we act according to beliefs about how to get it. When the

Humean theory is coupled with "reasons internalism," the idea that "you only have a moral (or normative) reason if you are capable of acting on it" (50), given that motivation is required for action, this results in a claim that contradicts the Kantian theory: one cannot have a "reason" to do what one does not want to do. In Smith's parlance, we would lose the objectivity of morality to maintain its practicality.[27] Hume's solution is for moral reasons to be contingent on desires but nevertheless universal because of the universality of sympathy, but this does not satisfy a Kantian conception of purely rational duty independent of desire, emotion, sentiment, or passion. The debate between moral reasons internalism and externalism thus centers on what "having a reason" or "there being a reason" implies about one's motivation, or lack thereof, to act on said reason, and in the background of this debate lies the Kantian versus Humean clash over the role of rationality versus emotion (or sentiment or passion) in understanding normative requirements.

Separate but linked to the question of how moral reasons relate to motivation, then, is the question of how emotions relate to motivation and if emotion specifically, and not just desire more broadly, is necessary for motivation. While desires are associated with feelings, desires understood broadly, in terms of direction of fit, do not necessarily have an affective dimension: it is possible to want something by having a disposition to pursue it without being aware of wanting it, let alone experiencing any specific feeling linked to that wanting in the moment. The concept of emotion is narrower. Tiberius describes contemporary theories of emotion as agreeing on their having four general

_____

[27] Smith's distinction between normative and motivating reasons is the central feature of his rejection of reasons internalism.

features: "importance (they are about something that matters to you), rationality (many emotions stand in rational relations to other psychological states and they can be evaluated for their appropriateness), feelings and phenomenology, and a tendency to cause action" (71). Different theories of emotion, such as the cognitivist account of emotions as judgments or beliefs or Jesse Prinz's (2004) Jamesian embodied appraisal theory of emotions as perceptions of bodily changes, are divided over how best to account for these features. However emotion is understood, sentimentalism, "the view that moral judgments essentially involve or are partly constituted by our emotions," makes sense of one interpretation of the practicality of morality: "as soon as we conceive of something as morally wrong, we are repelled by it, and as soon as we see something as morally admirable, we are drawn to it" (83). Such emotions, like judgments themselves, are seen as having "intentional content, that is, they are about or directed at something." Although Tiberius mentions "moods" on the same page, there is no discussion of affective states without intentional content. "All of the mental states" relevant to moral psychology, she emphasizes, have intentional content (9) – desires, beliefs, emotions – even if there are some complex cases like virtues, which might be interpreted variously as involving both directions of fit ("besires") or as dispositions that give rise to assorted intentional states.

Reflecting on this apparently exclusive focus on mental states with intentional content further helps to make sense of how motivation has been treated in moral psychology: as a conceptual gap to be filled when other concepts, like moral judgment or desire, are better understood. Presume that morality is necessarily (if not exclusively) concerned with conduct or action, an action necessarily has an object (whatever is being

81

done), and motivation is necessary for what is properly called action.[28] To ask what it means to be motivated to do some X is just to ask what adds up to motivation from among the intentional states of moral psychological concern: desires, beliefs, emotions, etcetera. All these issues related to motivation come downstream of the starting point, motivated action as an object of moral evaluation (if not the only one). This is entirely sensible. But this focus on action, given its centrality to morality and it arguably being the firmest evidence of motivation, has obscured the pervasive effect of habits and objectless moods, or what Matthew Ratcliffe calls "existential feelings" like feeling "alive" or "out of it," on motivation and, when moral psychology has taken up the topic of depression, contributed to the misinterpretation of depressive loss of motivation in terms of intentional states like desires. With this context, I turn to some authors who collectively present the strongest challenge to the Stocker-Smith account of depression: if depressive loss of motivation is just a loss of desires, then the experience should not be so painful.

2. Critiques of the Stocker-Smith Model of Depression

John Russell Roberts (2001) makes an analogy between depressed persons and persons acting out of compulsion: because both can be said to not "really" desire the actions (or lack thereof) that constitute their behavior, and because it is this conflict between their behavior and their "real" desires that explains their particular suffering or despair, neither group can be properly understood as experiencing a loss of desire. For

---

[28] Character would be understood in this context as comprised of dispositions to act, etc., and thus virtue ethics would be compatible with morality being necessarily but not exclusively concerned with action.

Roberts, though a person acting out of compulsion may seem to act out of desire, "they are alienated in some sense from the desire": they "fail to identify with it, in the sense that if asked they would say that they would rather not act on the basis of that desire." Similarly, a depressed person is unhappy with their own behavior; they "fail to identify with their deviant motivational states" (41) in the same way. It thus seems that depression "leaves one's evaluative outlook intact" such that depressed persons "suffer a great deal from their lack of ability to pursue their values," such as through intense feelings of guilt. A non-Humean who argues that beliefs may have intrinsic motivational character could identify the depressed person's "evaluative outlook" with their normative beliefs, but then would struggle to explain the depressed person's apparent loss of motivation absent any change in those beliefs. Roberts takes the Humean position that what it means to "really" want something is to value it and that valuing is a kind of desiring. If valuing is understood as desiring, then a depressed person is experiencing a conflict between their compulsive behavior and their desires, not a loss of desires at all. To "maintain the conceptual connection between desiring and valuing which the Humean account supplies, without facing the problem of explaining the depressives' lack of motivation," Roberts modifies what counts as desire satisfaction: we experience our desires as satisfied not just when we believe that our goal has been achieved but when our desire becomes "quiescent," when we *feel* satisfied in a way that seems to affirm our belief in our success; our feelings also offer a kind of information relevant to what we count as satisfaction.

Roberts points to how determinations about if a desire has been satisfied or not and predictions about if a desire will be satisfied or not involve "thinking with feeling":

83

our capacity to determine when we have done enough and may go on is in part affective, and it is that affective dimension of determining when a desire has been satisfied that is distorted in depression. The informational role of feelings in prediction is for Roberts the key to understanding a depressed person's inability to act. For a depressed person imagining possible courses of action, "nearly any generated hypothesis about action will meet with negative affect when tested" (54). And Roberts draws on empirical research suggesting that though depressed persons "may seem dormant" they are "internally very active" (57) in a manner inconsistent with a supposed lack of desire: though they tend to experience "a loss of confidence in their judgments of how to deal effectively with their environment," they also "show a tendency to be especially sensitive to information that may help them gain more control over a hostile world" (56). That is, depressed persons constantly try to predict what course of action will satisfy their desires, but every time, their ostensibly informative feelings predict failure, and over time, attributing this unrelenting felt impossibility of success to their own character, these predictions accrue into hopelessness. It is because depressed persons desire while constantly feeling that their desires are unattainable that they suffer. In this view, a depressed person unable to get out of bed from an overwhelming sense of futility is somewhat like someone anxiously checking the same lock on their door again and again: their desires or values or goals have not changed, but their capacity to know when their desires or values or goals have been satisfied is undermined by a persistent and compelling feeling that their desires remain, or will continue to remain, unmet.

I think that Roberts is right to think of depression in terms of compulsion, specifically compulsive negative evaluations of the possible outcomes of actions from

84

which follow apparent inaction, and to see the experienced symptoms of depression (like loss of interest or feelings of guilt and worthlessness) as emerging from these compulsive evaluations. I agree that the particular pain of depression, with its valences of emptiness or futility, stems precisely from the seeming impossibility of one's desires being satisfied, and thus that it is implausible to understand depressive loss of motivation in terms of a loss of desires. Although Roberts treats "feeling" broadly and does not distinguish between feelings and emotions like contemporary theory of emotion does, I think it can be useful when seeking to understand action or motivation to think in terms of feeling and not just emotion, as when reflecting on how "feeling threatened" might manifest through different emotions like fear, anger, or surprise. But I hesitate to equate the felt impossibility of action in depression with the uptake of bad information. Surely depressed persons do receive bad affective feedback when they imaginatively simulate action, and surely that feedback is forceful in a way that might drown out other predictive bases like more reliable perceptions and beliefs. The questions are, first, how the language of mental states is best applied to the idea of informative feelings and, second, if a depressed person's felt sense of impossibility or futility should be understood as a distinct and separable source of information. The answer to the first will depend on the answer to the second, that is, how these feelings work.

Roberts's conception of feelings as a distinct source of information is true in at least one important sense: a depressed person's sense of powerlessness can emerge from the awareness of their evaluative tendencies being out of step not only with those of other people but with their own perceptions and beliefs. I may believe that it is possible for me to get out of bed, to remember doing so countless times before and reasonably infer that I

85

may do so again, yet be crushed with dread at every attempt to imagine doing so. Not only is this tension confounding, but the seeming impotence of my reflectively endorsed beliefs to assuage that dread adds another layer of helplessness: the inability of thinking to lead to an action that seems futile makes the thinking seem futile too. Thus, understanding the experience of depression involves seeing it as leaving some perceptions and beliefs intact, since their intactness is what underscores their seeming inability to guide action despite having the same content as those mental states that have successfully guided action before. However, this story is complicated by the way in which, as Benedict Smith puts it, the "experience of depression involves changes to how the world is perceived and what the world offers a person by way of opportunities for action" (2013, 4). It may also be that the felt effects of depression are best understood as impacting at least some perceptions or beliefs.

Benedict Smith quotes from several authors who suggest not only that depression impacts beliefs but that the effects of depression on desires and beliefs are somehow intertwined in the feeling of one's tasks being impossible. Iain Law writes contra the Humean Stocker-Smith portrayal of depression "according to which one's belief profile remains unchanged whilst the independent motivating element has dropped away":

> My own experience of depression is not like this at all. In depression, everyday tasks take on an aspect of impossibility. I do not see, say, replying to an email, as an easy task that I am peculiarly lacking in motivation to perform. Rather, I see it as far more difficult than it actually is, or I see myself as not being up to the task … in my case desire has not come apart from belief. (2009, 354, quoted in Smith 2012, 4)

Law's account goes on to defend an Aristotelian conception of motivational cognitive states or "besires" and of virtue as knowledge.[29] Smith turns to the way in which depression impacts not just beliefs but habitual ways of perceiving opportunities for action: whereas seeing a laptop might usually "elicit a range of habitual actions" like "opening some emails messages and replying to them," a depressed person might see the laptop and experience "an explicit and distressing sense of alienation from these everyday activities; the opportunities that would otherwise be presented become closed-off or perhaps seen as impossibly difficult" (2012, 10). He turns to an autobiographical account that exemplifies this theme: "I'd look around the room and be almost scared by how solid the furniture seemed, how assertive the wallpaper. The folds in the curtain had authority" (Lewis 2006, 79, quoted in Smith 2012, 10). These quotes describe something like an aura, threatening and overwhelming, that enshrouds the experienced world and somehow affects desires, beliefs, feelings and emotions, habits, and character in one fell swoop.[30]

---

[29] Law argues that virtue is knowledge, or belief-like, and a depressed person, in lacking proper motivation because of altered beliefs, lacks virtue in a sense of having "one thought too many" compared to a truly virtuous person; however, this lack of virtue should not be morally condemned. I think that Law's argument is plausible but that it speaks poorly of virtue as a moral paradigm, especially for depressed persons, for a reason that I have not yet seen raised against virtue ethics: although virtue is the vocabulary of moral character, it is insufficiently fine-grained to improve moral character. To speak of a depressed person as "proud" or "cowardly" may be true as a description that predicts some of their behavior (if not moral matters involving altruism as per Cholbi) but will not be as useful for self-improvement as thinking in terms of how they perceive threats and anticipate rewards. If virtue ethics is concerned with the development of character, it seems to me it ought to move beyond the language of virtue. I also suspect that, if this objection has merit, blurring the moral and the prudential within the concept of virtue will only encourage a depressed person to moralize about the prudential. I will touch on virtue ethics again in Chapter 4.

[30] Unfortunately, the way that Benedict Smith employs the concept of embodiment, claiming that "the character of our bodily nature and not (or not just) our psychology is a basic feature of our agential capacities" (2), is flawed in a manner that presages Ratcliffe's problematic appeals to bodily feeling. Following Merleau-Ponty, Smith rightly points out that perception is structured by a body schema that inclines one toward habitual tasks such that no reflective mental representation of, e.g., a task as doable, mug as reachable, or what have you is necessary for action. But with the striking claim that motivation is "not psychological," Smith moves from the concept of affordances into a kind of direct realism: he seems to dispense with any mediating role for beliefs in motivation, locating motivation "in the world" and not "subjectivity" (9). This is what I call the phenomenological fallacy of a transparent body and which I will further discuss. In brief: "the body" is supposed to operate as a contingent version of a transcendental

The Roberts picture of depression as involving distinct and conflicting mental states and the Benedict Smith picture of depression as an all-encompassing shroud can be reconciled through Law's example of replying to an email in a way that helps to explain what I think of as the "infectiousness" of depression and anxiety. Law focuses on how, as soon as replying to the email is imagined, the prospect is suffused with dread: there is a singular encounter with the task as daunting. When I have this experience, I also remember that it has been easy and believe that it should be easy, leading me to try to reason with myself or talk myself down and, often, get increasingly frustrated as the part of me that seems in control of what I am able to do refuses to yield to argument. This can, as previously stated, produce distinct layers of feelings of powerlessness, both to act and to think in a way that affects action. Yet the experience also involves a kind of spiraling that comes from the way the dread creeps to every object of my attention. When I attend to the email, the email seems dreadful. When I attend to my rational belief that the email is not dreadful (or no more than any other of the many I have successfully written), that belief itself seems dreadful: its impotence is horrifying, a sign of my inadequacy, a whisper that maybe someday no thought I have will be able to move me at all and that I have no choice about when or if that day comes. When I attend to those second-order thoughts about my belief, *they* seem dreadful: just thinking them seems also to be evidence of my fundamental wrongness, maybe because the ability of others to accomplish such basic tasks suggests that I am the only one thinking this way, maybe

---

schema, the contours of which may be inferred from how we perceive rather than deduced purely rationally. But in an eagerness to downplay the role of reflection in identifying the structures of experience, embodiment can come to mean instead a *directly perceivable transcendental schema*, a mediator that itself needs no mediation to be grasped. Ratcliffe's existential feelings end up working just this way. My critique recalls Jacques Derrida's diagnosis of "haptocentrism," a dream of complete self-contact, in phenomenology from *On Touching—Jean-Luc Nancy* (2000).

because I suspect all this rumination is itself defensive or cowardly, meant to distract me from my obligations because I fear being held to account or never meant to keep my promises or knew all along that I was doomed to fail and hid this so that others would accept me. This sense of my own wrongness may be compounded by its epistemic and moral doubling: I am at once contradicting myself and failing to meet my obligations to another.

All these pitiful pseudo-explanations for the dreadfulness of each object of my attention come after the fact of the wave of sheer threateningness they send through me. That is, whatever I am paying attention to at any given moment – a thing in the world, my thought about that thing, my thought about my thought – is what will show up as overwhelmingly threatening, in whatever form that threat must take to seem like a threat. This "shroud" or "aura" aspect of depression or anxiety, the way they seem to touch or taint everything, refers not only to how any mundane or trivial task can come to seem to be somehow "too much" but how one's own mental states may be similarly affected just as they arise. Roberts is correct that a person in this state experiences a conflict between a feeling that failure is inevitable and a belief that success is possible. Law is correct both that the apprehension of the object of one's attention is straightaway wrapped up in that feeling and that, because of this, at least some of one's beliefs and desires do not simply "come apart" as in the Stocker-Smith account. If I am experiencing an aversion to a task and that is understood as a change in my desires, then it is linked to how I am intuiting what seems to me to be a fact about the world, that being the difficulty of the task, which would be understood as a change in my beliefs. It is hard to know how best to understand this "link" between desire and belief because of the way, as Law says, the "aspect of

89

impossibility" hits one all at once. The explanatory appeal of the concept of a "besire," a mental state with two directions of fit, makes sense in this context. But as my example of the spiraling of depression and anxiety suggests, that "aspect" may not be a self-standing intentional state but rather a kind of lens through which other states may be apprehended: what makes this quality lend itself to language like "shroud" or "aura" is not just that it happens to affect many objects of experience but that, when operative, it seems capable of affecting any intentional state one has about any object whatsoever.

I want to emphasize two possible conclusions from all this: first, that depression may not be properly understood if its effects are understood as limited to a single type of mental state, whether desires or beliefs or feelings; and second, given that depression involves loss of motivation, whatever depression affects may be an unidentified condition for the possibility of motivation. Given Roberts's argument about how depression causes pain, the Stocker-Smith account of depression as a loss of desires may be safely ruled out: depressive listlessness is best thought of not as a loss of desires but of a conflict between desires (or values or goals) and something else. In the framework of belief-desire psychology, that "something else" could be beliefs, possibly beliefs arrived at through the influence of the sorts of feelings that Roberts describes. Depression thus could be understood as primarily affecting beliefs and then having downstream effects on desires and habits – say, first believing that some course of action is futile and then becoming discouraged and losing interest – making sense of the holistic impact that Benedict Smith emphasizes without introducing any new categories to moral psychology.

No alternative account of depression that incorporates these insights into the conventional vocabulary of contemporary moral psychology has yet been developed, a

90

gap perhaps best exemplified by the interesting case of Steven Swartzer's 2015 and 2018

publications on depression and moral motivation. In "Humean Externalism and the

Argument from Depression," Swartzer challenges the Stocker-Smith portrayal of

depression on similar grounds to Roberts: being "unable to engage in activities that one

cares strongly about is part of what such situations are so heartbreaking" (2015, 9). But

Swartzer takes the idea "that depression sometimes leaves the agent's goals, concerns,

desires and commitments intact, but undermines their motivational efficacy" (8) in a

radical direction: he argues that depression raises the possibility that desires are not

necessarily motivational, just like beliefs, turning an example used by Michael Smith to

support Humean externalism instead into a challenge. In his later "A challenge for

Humean externalism," Swartzer develops this idea more generally:

> sometimes, we treat things as beyond our control when they really aren't. One
> way that this might happen is when sufficiently strong desires eliminate all
> available means from consideration. These possible courses of action are not
> considered live or serious options. When the only available means by which a
> given desire can be promoted are not live options, the agent might reasonably
> treat the object of desire as beyond her control, and thus might lack motivation to
> pursue it. This can happen without eradicating the desire, or making the agent
> want the desired end any less. (2018, 15)

However, the example of depression drops out of Swartzer's second paper and is

substituted with the case of amoralism to sharpen its challenge. A depressed agent,

Swartzer points out, might experience other "undermining motivational defeaters,"

whereas a hypothetical amoralist could be practically rational and still lack moral

motivation, setting a better stage for a debate over if such a figure should be understood

as lacking desire. Depressive loss of motivation may be neither good evidence for

Humean externalism nor good evidence against it. And when it no longer becomes clear

what a focus on depression might contribute to the preexisting framework of beliefs and desires, the discussion of depression drops out of moral psychology.

Setting aside Swartzer's provocative idea that even desire as Humeans understand it is not necessarily motivational, I will attempt to fill this gap and offer a parsimonious account of depressive loss of motivation that gathers the insights presented by these various challenges to the Stocker-Smith model of depression. By defining the effects of depression in terms of belief rather than desire, as occurring through the emergence of a compulsive "quasi-belief" in the impossibility of success, I produce an account compatible with belief-desire psychology that, like Swartzer's, neither offers evidence for or against Humean externalism.

3. A Substitute for Stocker-Smith: Depression as Quasi-Belief

Because a belief arising through compulsion might not be properly understood as a "belief" in the usual philosophical sense, I borrow the term "quasi-belief" from Robert Noggle's "Belief, quasi-belief, and obsessive-compulsive disorder" (2016). Noggle observes that the obsessive concerns of persons with OCD resemble beliefs in having the form of a proposition about some state of affairs that a compulsive action seems meant to change: in one of his examples, "Jane" repeatedly flips a light switch while thinking that it might get stuck in some in-between position that could cause a short-circuit and ignite. Yet even as she acts in a way that seems to be responding to this thought, she is also aware that the idea is far-fetched and, when she is not engaging in this compulsive behavior, she does not try to convince others of the dangers of light switches or otherwise act as though she thinks they are likely to catch fire. Jane thus seems to have a belief in

92

the sense that her mental state takes the form of a proposition that "drives emotion and motivation," but not in the sense that her mental state "is reported in the agent's sincere assertions" or "is used in practical and theoretical reasoning" (2). This point is sharpened by the example of persons with OCD who undergo exposure and response prevention (ERP) therapy to reduce OCD symptoms: the patients' "anxiety and compulsive motivation suggest that they believe the very thing that their willingness to undergo the process and resist the compulsive motivation suggests that they do not" (3). Such a person must be highly aware of the conflict between their obsessive concerns and their other beliefs, must have identified their obsessive concerns as irrational and want to be rid of them, and must understand that their obsessive concerns are so resistant to being reasoned away that exposure therapy, unpleasant though it must be, seems the best solution. Integral to this person's self-understanding and revealed by their struggle is a categorical difference between beliefs, with which a person identifies and which are at least minimally responsive to contrary evidence, and belief-like states like obsessive concerns, which persist even when disavowed and actively confronted.

Noggle dubs the latter "quasi-beliefs," mental states that are "belief-*like* in having propositional content and driving emotion and behavior in ways that would be rational if that content were true" but "lacking or severely deficient in the properties of affirmation and evidential responsiveness" (5). The property of affirmation claims that "functionally normal beliefs," or "the typical beliefs of psychologically normal adult humans," have two related qualities: someone with a belief is disposed to assert that belief as true, if only to themselves, and is generally disposed to use that belief as a premise in their theoretical and practical reasoning insofar as they "consciously attend" to said belief. The property

of evidential responsiveness states that "absent any strong attachment" to a belief being true, should someone be "simultaneously aware" of having a belief and that there is "strong, obvious, and direct evidence" that their belief is false, that belief "will typically disappear" (4). Noggle acknowledges that beliefs may be irrational while still qualifying as beliefs and does not elaborate on what constitutes consciously attending to beliefs or having a strong attachment to them. What these properties highlight is what is so striking about quasi-beliefs like obsessive concerns, that is, how completely someone with OCD willingly undergoing ERP has disidentified from and rejected propositions on which they nevertheless compulsively act. There is an asymmetry between the "agential authority" of beliefs and quasi-beliefs, what seems to oneself to be one's own "authentic voice," that is central to ERP and to a person seeking it out in the first place (9). Contrasting quasi-beliefs with similar ideas like Tamar Gendler's "aliefs," which do not have propositional content, Noggle emphasizes that the content of OCD-related anxiety involves complex propositions and the compulsive behavior driven by this anxiety can be sophisticated. Jane is specifically afraid that her house will burn down because a stuck light switch will start a fire and she is responding by repeatedly checking and flipping the switch. Her testing behavior would be rational if the probability assignments motivating her obsessive concerns reflected a real state of affairs.

If quasi-beliefs have propositional content, then there is a meaningful sense in which a person struggling with compulsive behavior is engaged in self-contradiction, and the concept of self-contradiction can make more concrete what may feel "impossible" about such situations from the perspective of one living them through. As Noggle clarifies, this is "not a matter of *the agent* asserting a contradiction. Rather, it is a matter

94

of the agent asserting something while her behavior and emotion are driven by a mental state that asserts something else" (13). A quasi-belief is an intrusive thought that assumes control and that the agent may thus be alienated from, finding themselves powerless to resist the force of assertions they do not themselves endorse. I would add to these observations that another element of this experience of powerlessness is the futility of reasoning through what appears to oneself as a contradiction. If my quasi-belief is that it is true that "the risk of this light switch burning my house down is extremely high" and my belief is that this is false, but my quasi-belief will not yield to evidence no matter how much supports my belief, I am left at a genuine impasse: there is no way to rationally proceed from this contradiction. All I can do is engage in ritualistic testing or hunker down and wait until the force of the assertion of my quasi-belief, and the feeling of being trapped in a contradiction from which there is no moving forward, ebbs. In ERT, repeated exposure to the fact that the belief is true and the quasi-belief is false is generating evidence, but the success of the therapy would not lie in the strengthening of the belief (because it is already sufficiently justified and convincing) but a subpersonal "wearing down" of the quasi-belief. In the absence of therapeutic support, quasi-beliefs may instead wear down the agent. Noggle notes that repeatedly experiencing oneself acting on quasi-beliefs that one believes are senseless is likely to diminish one's "ability to resist" (13): reasoning is fruitless when senselessness is no obstacle to behavior. I would add that if this experience is attended by a feeling of being stuck or trapped, this makes sense not only at the level of repeated behavior – that is, as arising from observing oneself being stuck in a loop – but at a mental level, of involuntarily violating a basic rule of

argument by compulsively asserting half of a contradiction and so being unable to usefully proceed in arguing with oneself.

This analysis of OCD meshes well with Roberts's account of depression in terms of compulsion and allows for a definition of depressive loss of motivation in terms of quasi-belief: specifically, depressive loss of motivation occurs through the emergence of a compulsive "quasi-belief" in the impossibility of success. If I am overwhelmed at the prospect of getting out of bed or dread replying to an email, I might understand my mental state as involving desiring to do so, believing that I can, and quasi-believing that I cannot. Both the persistence of my desire despite my inhibiting quasi-belief and the contradiction between my belief and quasi-belief serve to account for my suffering and multilayered sense of powerlessness. This gloss of the contents of my belief and quasi-belief as "I can"/"I cannot" might seem to be an oversimplification. But more complex and specific beliefs about these tasks – say, "I am too stupid to do this" versus "I am not stupid and therefore must be too lazy to do this" – can arise from thinking about, or attempting to explain or make sense of, a basic mental state that is best understood as a flat and unqualified assertion of impossibility. This account highlights how depressed persons may experience competing or contradictory beliefs that arise to "explain" or to rationalize that impossibility as somehow entangled, as all true even when collectively incoherent. The persistence of a basic quasi-belief in the impossibility of success even in the face of a contradicting belief in the possibility of success generates all sorts of possible explanatory beliefs, with any further contradictions that arise – like being both stupid and not stupid, thus lazy – seeming to blur into the background of the preceding contradiction around which all such thoughts orbit. In the context of moral psychology,

96

such an account is compatible with both externalism and internalism: it does not matter if desire is necessary for motivation or if there can be motivating beliefs so long as there is a quasi-belief in the impossibility of success preventing action. The quasi-belief theory successfully incorporates the objections and various proposed alternatives to the Stocker-Smith account of depressive loss of motivation and, in doing so, implies that the example of depression does not have much to add to the moral psychological debate with which it has been thus far associated.

I will call this adaptation of Noggle's concept of a quasi-belief from the context of OCD to the example of depression a "second-order quasi-belief." Noggle refers to quasi-beliefs that seem to have propositional content about a specific object, as with the person who repeatedly flips a light switch for fear of fire but is not equivalently fixated on any other potentially flammable items in their house. I define second-order quasi-beliefs that are directed not at external objects but at one's own mental states. This may produce what I have called the "shrouding" quality of depression, which operates in two ways: first, by "draping over" new intentional states as they arise, and second, in giving the depressed person the sense that these states are linked or entangled, e.g., by blurring one's ability to tell if one has lost a desire or changed a belief. To say that depressive loss of motivation is defined by a quasi-belief in the impossibility of success implies that this quasi-belief must be directed at one's own intentional states rather than objects in the world. That is, as soon as a desire or goal-relevant belief directed at the world arises, a quasi-belief directed at them asserts that successfully fulfilling that desire or acting on that belief will be impossible. The concept of a second-order quasi-belief captures the differences between Noggle's examples of OCD and my examples of depression while incorporating

97

criticisms of the Stocker-Smith model. Depressive loss of motivation has been successfully explained in terms of an intentional state, but a state directed at other intentional states, not at the world. Thus, the quasi-belief model represents what I take to be the most compelling way of making sense of demoralization that restricts itself to the terms of belief-desire psychology, which is to say, intentional states.

Despite the merits of this model, however, a fuller examination of the phenomenology of depression, thus far absent from mainstream moral psychological debates, exposes a questionable presumption that is operative in the focus on judgment: that the relevant conditions of possible motivation are always intentional states. Desires, beliefs, and emotions are all relevant to the possibility of motivation and all have intentional content. Depression can involve a changed relation to those same desires, beliefs, and emotions such that they seem empty and lose their motivational quality. And newly arising desires, beliefs, and emotions can be subject to the same emptying effect, suggesting that whatever is changing this relation is not directed at any specific mental states. What makes depression or anxiety seem like a shroud or aura is that it is not about anything and so could, effectively, become about anything. Understanding this aspect of depression helps to capture another facet of the sense of powerlessness or being trapped it may induce. Introducing the concept of a second-order quasi-belief accounts for that aspect while maintaining the belief-desire framework of moral psychology (so long as a quasi-belief is accepted as a kind of belief).

Yet we have also pushed that framework's limits to the breaking point. A second-order quasi-belief is a belief in the ways that matter philosophically, which is to say, technically: it is an assertion, albeit compulsive, about the world, albeit one's own mental

98

states. Nevertheless, it is difficult to distinguish a "belief" at this level of abstraction from a habit, which does not have objects about which a claim is asserted so much as triggers that activate a pattern of behavior. Habits, like quasi-beliefs, are intransigent: changes in one's beliefs are not necessarily or immediately reflected in one's habits, which is a common challenge in any attempt to modify behavior and is especially acute when an element of compulsion serves to entrench the habit. Consider an alternative explanation for depressive loss of motivation in terms of a habit, as a kind of reflex with no object, here framed as a pattern of feeling: "whenever a desire or desire-relevant belief arises, feel threatened." Granted that this presumes that the depressed person simply feels threatened and is only then driven to connect this sense of threat with some object in the world. Is depressive loss of motivation best understood as a compulsive "claim" "about" one's own intentional states when the analogy to conventional belief is so loose that this ostensibly intentional state is functionally indistinguishable from a non-intentional or at least intentionally ambiguous habit? Even if a second-order quasi-belief explanatory model *can* make sense of depressive loss of motivation, this functional equivalence suggests that explanation in terms of intentional states is the product of theoretical inertia: moral psychology is focused on intentional states as the relevant conditions of the possibility of motivation. But habits, for example, might also operate as such conditions.

In the next chapter, I argue that Matthew Ratcliffe's concept of "pre-intentional framing" attempts to capture this category of mental states, that is, those that modify other intentional states such that they could be understood either as 1) intentional states directed toward other intentional states or 2) intentionally ambiguous habits or dispositions that happen to interfere with intentional states. I turn to his phenomenology

99

of depression with three goals: 1) to enrich the concept of the shrouding effect that distinguishes depression from other conditions involving quasi-beliefs like OCD, 2) to evaluate Ratcliffe's concept of the pre-intentional after considering criticism, and 3) to subsequently modify it to meet likely objections were existential feelings to be introduced into moral psychology as a pre-intentional condition of motivation.

I then pivot to a discussion of John Dewey's theory of habits with three goals: 1) to develop the habit model of demoralization based on his analysis, 2) to determine if or how Dewey's theory should be modified based on debates between classical and neopragmatists, and 3) to show how this debate, perhaps due to the shared intellectual lineage of pragmatism and phenomenology, forecasts the epistemological problems faced by Ratcliffe in his use of the concept of existential feelings. I am ultimately critical of how Ratcliffe defines existential feelings and, based on Jussi Saarinen's criticisms and Ratcliffe's inadequate responses, I recommend a distinction between existential dispositions, which pre-intentionally structure intentional states, and existential feelings, which are epistemically valuable ways of becoming aware of the presence of existential dispositions. Even so, Ratcliffe's research helps to show what the example of depression has to contribute to moral psychology: not ammunition in the debate over moral judgment but a challenge for moral psychologists to consider non-intentional or intentionally ambiguous conditions of moral motivation.

In defining depressive loss of motivation in terms of the pre-intentional and modifying the category of the pre-intentional to encompass intentional states like second-order quasi-beliefs and intentionally ambiguous states like habits of feeling, I hope to do what I think Ratcliffe cannot: explain and justify prescriptions for existential hope as a

counter to existential guilt. That is, a depressed person may respond to existential guilt with indirect methods like treatment for compulsive quasi-beliefs or counter-habits to subvert compulsive habits, and in so doing come to better experience existential hope. Even if such a response is predicated on some preexisting hope, it may plausibly and even predictably lead to a depressed person becoming more hopeful.

I think that without my theoretical framing of pre-intentional states, this indirect response to existential problems can feel humiliating: "Great," a depressed person may think, "I've baby-talked myself out of an existential crisis again, and I have no better reasons for going on than I did before." What my refutation of the Stocker-Smith model ultimately aims to show is that a depressed person already has those reasons for going on: all their desires and beliefs are basically left intact by the presence of quasi-beliefs or habits of feeling, just rendered ineffective in leading to action. Now, a depressed person's experience of crisis or normative breakdown is revealing of a fundamental vulnerability: no desire or belief in itself is necessarily good enough for going on if it can be undercut by a pre-intentional state. It is reasonable to want to respond to the recognition of this vulnerability by developing a greater clarity of purpose.

But developing a greater clarity of purpose generally involves deliberation, and deliberation is difficult in the presence of a deliberation-defeating quasi-belief. The answer is to counteract the quasi-belief with an indirect method first and then, with the memory of one's vulnerability fresh, deliberate on one's purposes and the motives behind them. There is nothing shameful in starting with whatever "baby-talk," whatever pleasant or soothing distractions, one requires to defuse a downward spiral. There is no point in forcing oneself to give reasons in the face of a compulsion that is resistant to refutation

101

by reasons. What matters is that a depressed person does not forget their vulnerability once their episode of depression eases. If they are predisposed to have another such episode and are aware of this, they are not likely to forget.

When Solomon hung up on his friend before ascending to his rooftop, he was not at fault. It is important to say that a depressed person is not to blame for unvirtuous acts arising from depression. Solomon's act is a moral world apart from a cruel act like someone feigning depression for sympathy or to manipulate others toward greater devotion.

Generous readers may be nodding their heads at all this. But a depressed person may be thinking: "Why not? Did the wrong not originate from his existence? Must we commit ourselves to the weightlessness of consequences to exonerate the depressed? Is this attempt at exoneration supposed to convince me that I am not a burden on others, that I do not tire or hurt them, when I know that I am and do?" I think that this downward, self-abusing spiral amounts to one question: "Do I truly, in my heart of hearts, want to be kind?" The fact of the spiral itself is the answer: yes. Of course, reading "depressed persons are altruistic" will not counter existential guilt, because this guilt is not responsive to reasons, even good ones. Existential guilt will still demand that the existentially guilty justify their existence against an absolute yet totally vague measure. I think that this guilt must thus be countered indirectly. But, once it has been countered, it will not vanish. It will continue to generate motives and desires that involve questioning and doubt. It is the countering of existential guilt, through which it persists, that liberates depressed persons to express the care that, as we will eventually see, is itself a component of that guilt.

I continue to the next chapter, to discuss the central category of the way of ethical

thinking that best responds to existential guilt: that of the pre-intentional.

CHAPTER III: THE PRE-INTENTIONAL

Ratcliffe's account of existential feelings offers a rich depiction of the shrouding quality of depression and serves as the primary example of a pre-intentional framing that affects the motivational character of intentional states. In one striking example, he describes how the seeming inescapability of existential guilt intensifies a temporal dislocation in which time "slows down" as change seems impossible or irrelevant yet also "speeds up" as the sense of being threatened intensifies, amounting to a felt collapse of the future into a present burdened by an inescapable past (2015, 188). His fine-grained analysis of experience in depression is evocative of a line from John Dewey (1957): "the world is hopeless to one without hope" (154). Whereas in the previous chapter I developed a quasi-belief model of demoralization for which all relevant mental states are intentional, here I develop the alternative habit model of demoralization by drawing on Dewey's theory of habit as active. I propose that quasi-belief and habit should not be thought of as competing explanatory models accounting for the introspectively opaque causes of motivational loss. Rather, inspired by Ratcliffe's concept of the "pre-intentional," I will argue that quasi-belief and habit are two poles of one category: pre-intentional conditions of motivation are states that *might* be intentional states directed at other intentional states, like quasi-beliefs, or intentionally ambiguous states with effects on intentional states, like habits. This wide account of the pre-intentional is the one most useful, I think, for understanding depressive loss of motivation. But it diverges sharply from Ratcliffe's own account.

Ratcliffe's otherwise valuable phenomenology relies upon a flawed account of existential feelings as pre-intentional states, as critiques made by Jussi Saarinen show,

104

and his responses to these critiques are thus far inadequate. Ratcliffe's central error is an example of what I call a phenomenological fallacy of the transparent body: instead of treating the body as a contingent transcendental schema that is always operative but only potentially apprehended through phenomenological analysis, the "bodily" status of existential feelings serves as a justification to treat this transcendental schema as *itself constantly perceived*, conflating a mode of perception (bodily feeling) with what is perceived (the "mood" or "atmosphere" as grasped through or interpreted by one's dispositions). In identifying this fallacy, I follow Jacques Derrida's (2005) critique of Maurice Merleau-Ponty and what he calls the "haptocentrism" of phenomenology in *On Touching—Jean-Luc Nancy*: the identification of perceiving with what is perceived, even if equivocal as in the example of the "touching touch," serves as a form of absolute self-contact that ignores the extent to which that which structures our thinking or perception is inaccessible to us except indirectly via traces left behind. To resolve this problem and better translate Ratcliffe's insights into the context of moral psychology, I will argue that existential feelings are part of a feeling-disposition complex: rather than being pre-intentional frames themselves, they bring to our awareness the dispositions that are pre-intentional frames. For example, "feeling alive" makes me aware that I am disposed to perceive actions as possible and have motivational reserves while "feeling deadened" makes me aware of the inverse case. As I will show, this distinction alleviates some of the issues raised by a close examination of the concept of existential feelings, issues that would surely arise in attempting to introduce the pre-intentional into a belief-desire framework.

Before I proceed, I want to offer a speculative and (I think) sympathetic explanation for why I think Ratcliffe has erroneously identified feelings with the dispositions they bring to our awareness: phenomenologically, this distinction is not likely to appear. At times, from my own perspective, it may appear to be *because* I feel alive that I can act and *because* I feel deadened that I cannot. Existential feelings can seem to be a cause even when they are better understood as a symptom. Feelings *can* themselves be obstructions: an overwhelming feeling could "take up mental bandwidth," preoccupying a subject by its intensity or duration. Moreover, it would be a mistake to disregard existential feelings in favor of a focus on existential dispositions: it can be useful to see a feeling itself as not "merely" a symptom, as something to itself be coped with or managed rather than as just a sign of some other state needing attention. For example, a mindfulness practice might encourage association with rather than dissociation from one's overwhelming feelings and, in the process of experiencing the extension of time and the possibility of staying with those feelings, succeed in shifting one's pre-intentional framing. All this is to say that introducing a distinction between existential feelings and existential dispositions is not to deemphasize existential feelings as theoretical or practical matters of concern, only to clarify them as such.

Distinguishing between existential feelings and existential dispositions not only improves on Ratcliffe's theory but accomplishes a goal in line with my broader project: in accounting for how feelings can be distinguished from dispositions practically, meaning as they are experienced, it gives argument or reasoning an indirect therapeutic role. Consider the case of talking oneself down from a surge of anxiety. If I am experiencing an overpowering and compulsive sense of being threatened, I will not be

106

able to reason with myself conventionally: that feeling (or quasi-belief) will not be responsive to evidence like memories of successfully overcoming adversity or reasonable predictions about the limited consequences of failure. But if I can, say, ask myself what I am threatened by and how threatening I think it is, giving myself time to do so, I will start to notice the gap between my beliefs and quasi-beliefs. As one part of me insists implacably I am in crisis, another, previously suppressed by that sense of crisis, will object. As I notice this gap, so I may notice that my feeling of panic involves a distinguishable predisposition to evaluate imagined actions or responses as futile, sharpening my sense that what may feel impossible are acts I believe and know to be possible.

Two aspects of this process are valuable. First, it involves a similar decompression of time as in the mindfulness example above, even if it is more reflective than mindful. Second and distinctly, it involves experiencing oneself as rational just as one is suffering from a loss of agency. I think that there is value not just in being rational, being able to endorse or discard beliefs, but in being able to feel oneself to be rational, to be assured that one retains that capacity even when quasi-beliefs resist it. A contradiction between one's beliefs and quasi-beliefs, made irresolvable by the evidence-unresponsive nature of quasi-beliefs, may seem to expose the impotence of one's own rationality, but one need not be able to argue oneself into submission to feel oneself to be an agent; sometimes, one need only be able to name what is occurring. What I aim to identify here is a way in which theory or argument, even when it will fail at the theoretical purposes of ascertain true beliefs or changing false beliefs – not just is likely to fail, but is guaranteed to fail by the nature of the epistemic states involved! – can succeed at a practical purpose

of easing a sense of alienation and becoming able to go on. Thus philosophy, understood in an orthodox manner, takes on an unorthodox purpose.

I now turn to summarize Ratcliffe's phenomenology of depression as it pertains to the shroud effect and the concept of pre-intentional framing. After offering his account, with a focus on the themes of existential guilt and hope, I pause to delve into Dewey's theory of habit for three reasons: 1) to apply his theory to my habit model of demoralization and show how habit may be understood as pre-intentional rather than merely reflexive and strictly non-intentional; 2) to show how the common assumptions of Deweyan pragmatism and Husserlian phenomenology, rooted in the work of William James, may produce confusion around the epistemic function of the experience of context, e.g., moods or situations; and 3) to suggest that Dewey and Ratcliffe share a normative concern for elevating stigmatized concepts that can lead to the overstatement of their descriptive claims. Having done so, I return to Ratcliffe's phenomenology and consider Jussi Saarinen's 2018 critiques of existential feeling and Ratcliffe's 2020 response, support Saarinen's critiques, and detail my proposed distinction between existential feelings and existential dispositions. I redefine Ratcliffe's category of the pre-intentional to refer to mental states that modify other intentional states and that, being opaque to introspection, could be understood either as 1) intentional states directed toward other intentional states, like quasi-beliefs, or 2) non-intentional dispositions that happen to interfere with intentional states, like habits in the Deweyan sense. This redefined category is, unlike Ratcliffe's original version, compatible with belief-desire psychology while offering a new and useful concept for understanding the variety of possible conditions of motivation and causes of demoralization. I hope to show also that

108

Ratcliffe's and Dewey's normative goals alike are better served with the introduction of sharper distinctions between existential feelings and existential dispositions and between qualitatively unified and qualitatively ambiguous situations, respectively. The needs of depressed persons in coping with compulsive reason-seeking are better met when the advantages and disadvantages of habits and feelings – their efficiency and potency, their intransigence and fallibility – are more clearly defined. And a philosophy with an explicitly normative overarching goal of theorizing a more hospitable world is compatible with, and I hope better served by, clear definitions.

1. Existential Feelings and Pre-Intentional Framing

For Ratcliffe, depressed persons undergo an "existential change," an "all-enveloping shift in one's sense of 'belonging to a shared world'" (2015, 14). Key to this notion, developed from phenomenologists like Husserl and Heidegger, is that all persons have a phenomenological sense of a world, "a backdrop against which we are able to adopt attitudes of whatever kind of whatever kind towards states of affairs within the world," a sense of one's presence and the world's reality that is irreducible to a set of judgments or beliefs and takes on a pervasive background quality of enabling, of I-can. Ratcliffe argues that in depression the phenomenological world or "sense of reality and belonging" is "not completely lost" but is "profoundly *altered*." What changes in depression is not a specific state of perceiving, feeling, believing, or remembering but "a change in the *structure* of perceiving, feeling, believing, and remembering, attributable to a disturbance of 'world'" (19). This structure is what Ratcliffe calls the "pre-intentional," that which modifies intentional states without itself being an intentional state. By

explaining the effects of depression with reference to this pre-intentional structure, Ratcliffe aims to show the unity of depression as it is experienced and so respond to the etiological challenge of defining mental illnesses: whereas many illnesses are defined by their cause, say a bacteria or virus, mental illnesses are defined as collections of symptoms that are not all necessarily present, making diagnosis more challenging or dubious. If it is difficult to pin down what depression's symptoms have to do with one another because its causes are so complex, then it might be easier to understand depression as a whole in terms of how it modifies one's experiences as a whole, through appeal to phenomenology.

Ratcliffe introduces the term "existential feelings," distinct from emotions or moods conventionally understood, to describe how we "find ourselves in a world." In *Feelings of Being* (2008) he offers an illustrative list:

> People sometimes talk of feeling alive, dead, distant, detached, dislodged, estranged, isolated, otherworldly, indifferent to everything, overwhelmed, suffocated, cut off, lost, disconnected, out of sorts, not oneself, out of touch with things, out of it, not quite with it, separate, in harmony with things, at peace with things or part of things. There are references to feelings of unreality, heightened existence, surreality, familiarity, unfamiliarity, strangeness, isolation, emptiness, belonging, being at home in the world, being at one with things, significance, insignificance, and the list goes on. People also sometimes report that 'things just don't feel right,' 'I'm not with it today,' 'I just feel a bit removed from it all at the moment,' 'I feel out of it,' or 'it feels strange.' (68)

Ratcliffe's point is these feelings are distinct at the felt level: there is a thus far underexamined category of feeling that tracks our existential changes, that is, changes to the pre-intentional structure modifying how we experience perceiving, believing, remembering, feeling other emotions, and so forth. Furthermore, what unites these feelings is that they are feelings of "possibility," which implies that "we experience *possibilities*" at the level of feeling and not just, say, by reflectively inferring what is

110

possible for us, and that depression "involves a change in the *kinds* of possibility that are experienced" (42). Ratcliffe describes the varieties of possibility detectable through existential feeling, such as degrees of determinacy or uncertainty, sense of difficulty, significance (e.g., utility, safety, danger), and interpersonal accessibility (relation to self and others, e.g., for whom the possibility is possible). Most boldly, Ratcliffe seems to claim that existential feelings do not merely make us aware of, e.g., a set of dispositions to perceive in certain ways, etc., constituting our pre-intentional structures, but *are* those pre-intentional structures: "our access to kinds of possibility is itself integral to our experience (rather than being a non-phenomenological disposition to have certain kinds of experience)" (51). This is bold because it implies that we are always experiencing "our access to kinds of possibility," and if this access *is* the pre-intentional and that experience takes the form of a feeling, then wherever there is a pre-intentional structure, there is always a feeling of some sort.

To justify this, Ratcliffe emphasizes that existential feelings are "bodily feelings": "a bodily feeling can at the same time *be* a sense of the salient possibilities offered by a situation" (59). Here he draws on the Husserlian distinction between *Leib* and *Korpor*, the "noetic" living body through which one perceives versus the "noematic" body as object of perception. For example, if I desire to take a sip of tea from my mug, I do not need to reflectively infer that the mug is within reach (or remember a previous such inference) because the mug seems to present itself to me as "in reach," an aspect of perception quietly contingent on how accustomed I am to the length and functions of my arms and hands. Upon touching the mug, one object of my perceiving may become a sensation localized in my hand, this being a perception of my body as *Korpor*, but the

structural element of perceiving informed by my familiarity with my body as *Leib*, which facilitates cognitive shortcuts as in the above example, is not localized or objectified in the same way. Similarly, according to Ratcliffe, existential feelings are not necessarily localized in the body in the manner of dread seeming to dwell in the pit of one's stomach: they "consist in a diffuse, background sense of bodily dispositions" through which "the body provides a kind of orientation through which the world is ordinarily encountered in the style of confidence or certainty" (59). But Ratcliffe goes a step further still: he argues that both noematic and noetic feelings "are experienced against a backdrop of existential feeling" (83), a third aspect of the body's involvement in perception, "a general style of experiencing that determines the *kinds* of more localized feeling a person is able to experience, as well as the overall *balance* between noetic and noematic aspects of experience" (85). Saarinen summarizes the claim: "Ratcliffe identifies kinaesthesia, proprioceptive awareness, interoceptive awareness of visceral feelings, and tactile background feelings as the bodily ingredients of existential feelings … [that] interact with each other in a variety of ways to produce a unitary bodily sense that has its own phenomenology" (2018, 11). To reinforce his position that existential feelings are a transcendental structure of perceiving, etc., that is directly felt, Ratcliffe claims through emphasis on their "bodily" dimension that existential feelings are a kind of meta-structure of *Leib* with their own unity distinguishable through feeling.

By positioning existential feelings as felt transcendental structures, Ratcliffe becomes able to show how the pervasive influence of depression on experience – including experiences of "the body, the world, hope, guilt, agency, time, interpersonal relations, and/or certain kinds of belief" (2015, 250) – is *itself* experienced in a singular

112

way, as a complex enveloping feeling that both feels "enveloping" and actually envelops, as in constrains the scope of, mental states including other feelings or emotions. This double envelopment is what I have referred to as the shrouding quality of depression. Although I will criticize his use of the concept of existential feelings, I agree that there are such enveloping feelings and that attending to them through phenomenology yields insights into what is happening in depressive loss of motivation. I aim to convey the power and importance of this phenomenological approach to depression by turning to two of these examples particularly relevant to ethics: existential hope and existential guilt.

Existential hope is the felt backdrop and precondition of intentional hope, that is, hopes for specific outcomes. Because loss of existential hope and intentional hopes can be described similarly, Ratcliffe argues, "loss of hope in depression is often misconstrued in intentional terms, and the profundity of the experience is not acknowledged" (99). Depression may involve the loss of intentional hopes. But in the background of intentional hopes generally is a "specific kind of anticipatory structure that includes possibilities such as 'things could change for the better' and 'bad things might not happen'" (108) and the loss or change of this structure can be felt in a manner distinct from the loss or change of intentional hopes. In existential hopelessness, all anticipation takes place "in the guise of fear, dread, or horror," and because "the world offers only danger," the "experience frequently includes a sense of one's own impotence and, with it, an inability to depend on others." With this existential change comes the loss of "a structured system of significant possibilities, of the kind required for the intelligibility of projects, goals, and hopes" (112). Existential hope thus resembles what Jay Bernstein, following Annette Baier and Susan Brison, calls "trust in the world," defined as

113

the existential confidence that permits the rational suppression, overlooking, forgetting, or fortunate ignorance of each individual's utter dependence on surrounding others, and hence each's categorial helplessness; with our helplessness no longer in conscious view, we can attend to the world rather than ourselves, or ourselves as fully worldly beings. (2011, 398)

The possibility of hoping for anything depends on a diffuse and likely unarticulated conviction, more inertia than belief, that one is not powerless or forsaken. A person lacking existential hope can feel its absence in the specific phenomenological tenor of existential hopelessness's blend of impotence and isolation. To attempt to understand this phenomenon as the loss of a set of intentional hopes risks overlooking how existential hopelessness precludes intelligibly adopting new hopes and how existential hope thus serves as a precondition of intentional hope.

Similarly, though existential guilt can be misunderstood as intentional guilt, it is disruptive to the structure of intentional guilt, which has an object and so a starting point for conceiving of how the guilt could be alleviated. Ratcliffe describes a spectrum of guilt from most straightforwardly intentional to existential, one that maps onto the felt "depth" of guilt: 1) feeling guilty about something specific, 2) feeling guilty about something but not knowing what, 3) feeling that one really is guilty of something specific, 4) feeling "that one simply *is guilty*, guilty independent of any particular act or omission, as though there were a moral flaw in one's being" (133), and 5) feeling irrevocably guilty, as though one's seeming moral flaw were irredeemable. Though all five kinds of guilt may occur in depression, Ratcliffe characterizes 4 and 5 as existential guilt and frequent in depressed persons, symptoms that "cause considerable suffering and are sometimes singled out as the most troubling" (135). What is striking is that although 4 and 5 feel like guilt, they lack some of its core elements, those being a sense of whom one has wronged

114

and what act wronged them. In explaining how existential guilt feels like guilt in the absence of those elements, Ratcliffe emphasizes "certain common themes: a focus on past deeds, recognition that effects of one's deeds are unchangeable, an awareness of estrangement from others, a sense of having done wrong or of being intrinsically flawed, and an anticipation of being harmed or punished" (134). Existential guilt has the form of guilt, its structure and trappings, without its intentional content. Moreover, because of this, existential guilt supersedes intentional guilt: if one is essentially guilty, then the prospect of being forgiven for specific wrongs is empty. As existential hopelessness makes pursuing one's projects unintelligible to the one experiencing it, so existential guilt makes redressing one's wrongs unintelligible.

These examples show how the concept of existential feelings as a felt pre-intentional structure helps give voice to real experiences of depression that are otherwise difficult to distinguish from their intentional variants. This distinction between the existential and the intentional, akin to the Heideggerian distinction between the ontological and the ontic, may help with the understanding and self-understanding of depressed persons and so help to counteract the isolation that depression both presents and can cause. For the purposes of this chapter, these concrete examples also clarify the importance of two claims: that the pre-intentional is a useful category for understanding loss of motivation in depression and that pre-intentional structures, and changes to them, can be felt.

Ratcliffe's theory of existential feelings bears a striking resemblance to Dewey's theory of qualitative experience, which for him defines the "situation" that structures experience and contextualizes deliberation and inquiry. In turn, Dewey's theory of

inquiry emphasizes the positive role of habit as actively structuring patterns of thought and action and the subsequent importance of education as habit-shaping. Drawing on Dewey's theory of habit to develop the habit model of demoralization also offers the opportunity to explore the critiques of Dewey's understanding of qualitative experience that have arisen in debates between classical and neopragmatists. In so doing, I will be able to show how Ratcliffe's misuse of the concept of existential feelings based on a fallacy of the transparent body is an echo of a similar past descriptive misstep arising from laudable normative goals. This resonates with the central theme of the intrapragmatic debate: the place of normativity in philosophy.

2. Dewey on Habit and Qualitative Experience

Dewey's theory of habit is significant both as an alternative example of the pre-intentional and because of the epistemological questions its associated theory of qualitative experience raises which, as I will show, parallel some of the challenges faced by Ratcliffe's theory of existential feelings. I have argued that depressive demoralization could be attributed to the presence of a quasi-belief in the futility of action but, equally plausibly, could also be attributed to a habit or disposition to feel threatened, downstream of which are defensive paralysis and rumination. The equal plausibility of both models clarifies the scope of the category of the pre-intentional as applying to mental states that could be intentional (like quasi-beliefs) or nonintentional (like a reflex) and which introspection cannot further disambiguate. The usefulness of Dewey's theory of habits here lies in the deliberate ambiguity of its intentional status: habits are like background assumptions operative in action and perception, affecting the scope of what seems

possible or probable or relevant, and so neither have a straightforwardly intentional status like beliefs or desires (in belief-desire psychology) nor a straightforwardly nonintentional status like a mere reflex. This understanding of habits not only resonates with the notion of the pre-intentional as such but fits into a model of consciousness in which pre-intentional states are a regular condition of the possibility of motivation, as I am arguing is the case in demoralization.

But what is also illuminating is how habits fit into the broader framework of Dewey's philosophy: as the conceptual pivot point for a Darwinian-inspired pragmatic and anti-dualistic theory of inquiry that grounds anti-hierarchical and explicitly democratic philosophies of education and politics. Habits represent both the possibility and the challenges of radical change for the self and society. On the one hand, Dewey argues, as much of what is called fixed nature in living things evolved and can still evolve further, so too is the character of a person or people the product of entrenched habits that can change and be changed. On the other hand, habits are intransigent even in the face of great and willing effort. The goal of education is thus not just to change beliefs but to change habits to better realize those beliefs in everyday perception, thought, and action. In turn, collective deliberation in the political sphere is not just a method of arriving at a correct conclusion or mutually acceptable compromise but a continuing process of shaping habits of conscientiousness and open-mindedness. "When," Dewey writes in *Liberalism and Social Action* (1963), "I say that the first object of a renascent liberalism is education, I mean that its task is to aid in producing the habits of mind and character, the intellectual and moral patterns, that are somewhere near even with the actual movements of events" (61). And, as he develops later in *The Public and its*

117

Problems (2016), the awareness that "man acts from crudely intelligized emotion and from habit rather than from rational consideration" (183) risks undermining confidence in the legitimacy of democratic control and motivating a turn toward elitism. Running throughout Dewey's philosophy is an emphasis on the implicit operation of habit, and a vision of inquiry founded on the cultivation of problem-solving habits, as a potentially democratizing bridge between seemingly separate metaphysical and social categories. The mind may transcend the body, but patterns of thought are no less shaped by habit than physical activities are by practice; fine arts may be elevated above vulgar crafts, but both are equally for Dewey products of inquiry, and so while this elevation is revealing of social values, it is not based on any actual features of the things themselves. Thus are the justifications for esteeming the artist over the "mere" artisan or the manager over the "mere" laborer undermined at the epistemological level. Understanding the role of habit in inquiry leads to a broad rejection of sharp discontinuities between categories, a theoretical stance that supports an egalitarian worldview.

This worldview, however, sharpens a central problem in the scholarship on Deweyan metaphysics: the role of qualitative experience in defining the "situation" to which inquiry, and the habits it encompasses, is said to respond. As habits operate in the background of thought and action, so the situation is the background against which any object appears: it is, for example, in a tense situation that threats emerge. Dewey understands the scope of a situation as defined by its "pervasive and unifying quality," the tenor or tone of the experience. But it follows that persons having different qualitative experiences are in different situations, which is in tension with a notion of inquiry as collective deliberation or problem-solving if the participants' qualitative experiences

118

conflict. Here, Dewey's epistemic and political positions may be at odds, which is clearly

not his intention. The best way to preserve Dewey's democratic view of inquiry, I think,

is to deemphasize the role of qualitative experience in setting the bounds of a situation, as

it is not just plausible but likely that two persons in a "problematic situation" together

will have wildly different qualitative experiences; after all, this might itself be the source

of a conflict.

I propose 1) a distinction between the qualitative aspects of situations, defined by

a mood, and the non-qualitative aspects of situations, defined by historical and material

factors, and 2) another between a problematic situation defined by a pervasive unifying

quality and one defined by qualitative ambiguity or centered on a qualitatively opaque

problem. I also explain why Dewey would endorse this modification. His emphasis on

qualitative experience strikes me as an overcorrecting reaction to what he sees as an

overintellectualized and elitist view of inquiry, according to which deliberation is

primarily or solely a matter of the arguments and abstractions that are foregrounded and

thus a matter for experts rather than a matter of the background habits and practices that

facilitate, and are through reenactment reinforced by, that foreground. I make this detour

to show how Dewey's epistemology risks working against itself and his politics when too

much rides philosophically on one's qualitative experience.

This risk will prove instructive in the case of Ratcliffe's theory of existential

feelings because he similarly assigns them too much self-standing significance compared

to other mental states, even more so than Dewey does with qualitative experience, and so

renders their epistemic function incoherent. This is no coincidence, as the concept of

existential feelings is descended from the same intellectual lineage: Dewey's "situation"

is a version of William James's "fringes," which Edmund Husserl develops into the phenomenological "horizon" operative in Martin Heidegger's variations on "being-in-the-world," such as his understanding of "mood," on which Ratcliffe draws.[31] The core insight, the phenomenological variation on the Kantian transcendental, is that some "background" is always operative in a structural or conditioning capacity behind whatever is "foregrounded" in one's attention. Existential feelings are for Ratcliffe what qualitative experience, and the situation it defines, is for Dewey. But as I will show, Dewey's theory of inquiry acknowledges that qualitative experience is just the beginning of a process of problem resolution that involves conceptualizing quality. Ratcliffe, committed to the primordiality and ubiquity of existential feeling, instead insists that existential feelings cannot be properly conceptualized, doubling down on framing them as transcendentally conditioning and unconditioned by other mental states. Since Dewey's theory of habits will also be relevant to the discussion of values and norms in Chapter 5, I take this opportunity to expand on it and the surrounding debate to support my argument in this chapter for a modification to Ratcliffe's theory of existential feelings.

Dewey's theory of habit is a key component of a gradualist philosophy that finds its broadest expressions in his metaphysics of the continuity of experience and nature and in his corresponding logical principle of continuity. "Experience" is for Dewey (1958),

---

[31] Dewey acknowledges his debt to William James, if critically, in a footnote in "Qualitative Thought" (1930): "The 'fringe' of James seems to me to be a somewhat unfortunate way of expressing the role of the underlying qualitative character that constitutes a situation—unfortunate because the metaphor tends to treat it as an additional element instead of an all-pervasive influence in determining other contents." For more on the historical connection between Husserl and James, see Saulius Geniusas's (2011) "William James and Edmund Husserl on the Horizontality of Experience" and its sources. Ratcliffe (2015) cites earlier phenomenologists such as Husserl and Heidegger in the introduction to *Experiences of Depression* (e.g., 2).

like James's "blooming, buzzing confusion," a unified whole: "it recognizes in its primary integrity no division between act and material, subject and object, but contains them both in an unanalyzed totality" (8). "Nature" refers to the states of affairs that give rise to experience, "an affair *of* affairs, wherein each one, no matter how linked up it may be with others, has its *own* quality" (97). Inquiry is a matter of reckoning with the world as it is already thrust upon us, an ongoing interaction between organism and environment rather than an active mind attempting to apprehend inert matter. Dewey's metaphysics offers a picture of the world compatible with his position on epistemology: "the world must actually be such as to generate ignorance and inquiry" (69). What follows, as he explains in his *Logic: The Theory of Inquiry* (1938), is a rejection of appeal to epistemic grounds external to or transcending the world as experienced:

> The primary postulate of a naturalistic theory of logic is continuity of the lower (less complex) and the higher (more complex) activities and forms … it precludes reduction of the 'higher' to the 'lower' just as it precludes complete breaks and gaps. The growth and development of any living organism from seed to maturity illustrates the meaning of continuity. … What *is* excluded by the postulate of continuity is the appearance upon the scene of a totally new outside force as a cause of changes that occur. (23-24)

This naturalistic worldview defines the core of Deweyan epistemology: no "totally new outside force," whether divine revelation or indubitable intuition, may verify or validate a claim that brings inquiry to an end, which leaves only practices of perpetual revision and imperfect forms of justification. For Dewey, this "continuity" between experience and nature or organism and environment is a virtuous circle, "not vicious" because "it is existential and historic," a naturalized dialectic of both individual and communal inquiry (1960, 247). Successful inquiry results in an organism's increased control of its environment, as well as "enriched meaning and value in things, clarification, increased

121

depth and continuity—a result even more precious than is the added power of control"

(1958, 11). By focusing on practical problem-solving as the paradigm case of knowledge-

seeking, Dewey aims to both describe inquiry as he thinks it actually is, a common

practice rather than the province of elite specialists, and to make the normative case for

what it could become, a more accessible and effective way of life guided by the

development of "intelligent" habits and not reliant on any special powers or faculties.

Habit is the crux of inquiry for Dewey in both its success and failure: habits

actively shape ideas and impressions, operating in the background to direct thought and

perception, but when circumstances change or impulses arise such that habits produce

conflict, a problematic situation emerges that begins inquiry anew. In *Human Nature and*

Conduct (1957), he declares that "an idea gets shape and consistency only when it has a

habit back of it" (30) and even that "were it not for the continued operation of all habits

in every act, no such thing as character could exist," for "character is the interpenetration

of habits" (37). He explains why he places so much significance on such a modest notion

in a passage worth quoting in full:

> The word habit may seem twisted somewhat from its customary use when
> employed as we have been using it. But we need a word to express that kind of
> human activity which is influenced by prior activity and in that sense acquired;
> which contains within itself a certain ordering or systematization of minor
> elements of action; which is projective, dynamic in quality, ready for overt
> manifestation; and which is operative in some subdued subordinate form even
> when not obviously dominating activity. Habit even in its ordinary usage comes
> nearer to denoting these facts than any other word. If the facts are recognized we
> may also use the words attitude and disposition. But unless we have first made
> clear to ourselves the facts which have been set forth under the name of habit,
> these words are more likely to be misleading than is the word habit. For the latter
> conveys explicitly the sense of operativeness, actuality. Attitude and, as ordinarily
> used, disposition suggest something latent, potential, something which requires a
> positive stimulus outside themselves to become active. If we perceive that they
> denote positive forms of action which are released merely through removal of

some counteracting "inhibitory" tendency, and then become overt, we may employ them instead…. (39-40)

Dewey is insistent that habits are always active for, I think, two main reasons. First, habits include acquired patterns of perception and prospective action-preparation that, rather than lying dormant until stimulated like a reflex, act on stimuli in the sense of filtering how they are presented to one's consciousness or call on one's attention. Second, Dewey is sensitive to the denigrated status among intellectuals of "second terms" like passivity, and his theoretical writing is often driven by a fierce demand that the lesser of two terms in any duality – the one that is typically thought to be merely parasitic on, derivative of, or acted on by the other – be understood as productive in its own way. This second point is especially relevant because it may contextualize some of Dewey's more sweeping claims. When Dewey is offering a descriptive account, he is typically also at the same time engaged in a normative project of elevating what he thinks other intellectuals take to be insignificant or unworthy of study. He may, in pursuit of that goal, sometimes overstate his case in a manner that his philosophical system does not require.

The process of inquiry culminates not just in a resolved problem, signaled by a consummatory experience, but in modified habits, a point that is relevant equally to Dewey's epistemology, ethics, and aesthetics. "Conflict of habits," or of habits with a changing environment, "releases impulsive activities which in their manifestation require a modification of habit, of custom and convention" (81). In this context, Dewey's understanding of deliberation is analogous to how a child plays: it is "a dramatic rehearsal (in imagination) of various competing possible lines of action" that "starts from the blocking of efficient overt action, due to that conflict of prior habit and newly released impulse" and during which "each habit, each impulse, involved in the temporary

suspense of overt action takes its turn in being tried out" (179). Deliberation is done well just insofar as it allows for the greatest adaptability to the widest range of circumstances, or as Dewey puts it, it "is rational in the degree in which forethought flexibly remarks old aims and habits, institutes perception and love of new ends and acts" (186). When Dewey refers to "intelligent" habits, it is this definition of rationality as practical effectiveness that he has in mind: "intelligence is concerned with foreseeing the future so that action may have order and direction" (220). The remaking of habit is intelligent when it increases adaptability and, significantly for his ethics and aesthetics, the more adaptable one is, the wider an array of circumstances may be experienced as meaningful, and thus the more meaningful one's life may seem to be. "Morals means growth of conduct in meaning" (259), and one is growing in when one endeavors "to convert strife into harmony, monotony into a variegated scene, and limitation into expansion" (260). Dewey's continuity of experience and nature arguably reaches its fullest expression here, where his epistemology meets his ethics and aesthetics: the pursuit of knowledge is an organism acting to respond to and shape its environment, its external world, and in so doing it is gradually rewarded with increasingly enriched experiences. No transcendent or ultimate purposes are required to motivate or justify a journey every sentient being undertakes.

Habit's mediating function is such that it effectively serves as the "continuity" in the "continuity of experience and nature," and the concept shows up at every stage of Dewey's philosophy from logic to ethics. He thereby inverts the hierarchy of foreground over background, both at the level of in-the-moment experience and social esteem: changing habits to become more adaptable and enriching might even be said to be the

124

real payoff of inquiry, not just a side effect of problem-solving, insofar as such change is what shapes and potentially enriches an organism's experiences going forward. But if this is so, more must be said about the "problematic situation" in which habits conflict and inquiry is renewed. Where Dewey sought to theorize a naturalistic metaphysics that would reflect his egalitarian politics, later pragmatist critics will challenge the terms of that metaphysics and even its necessity. Before taking Dewey's robust theory of habit as a model for the pre-intentional, I will clarify to what extent it is dependent on Dewey's contentious understanding of the situation and qualitative experience.

Dewey's definition of the situation is guided by his emphasis on the active operation of background on foreground and, in complementary fashion, his rejection of atomism both metaphysically and politically. In his essay "Qualitative Thought," he defines a situation as a "complex whole" that "is held together in spite of its internal complexity by the fact that it is dominated and characterized throughout by a single quality," in which any object "is defined in abstraction from the whole of which it is a distinction" (1930, 197). Though objects *of* our attention may be foregrounded *in* our attention, their salience and significance are determined contextually, by the background situation. As James Gouinlock (1972) summarizes the point: "the quality of the situation, or context, brings about the kind of behavior which will effect a transition to a particular determination in which objects or relations are discriminated, and discriminated in a way which is stimulated by the original quality" (8). This is consistent with Dewey's broader concerns and claims, but he specifies in this essay that meaning, both experienced and linguistic, is qualitative. He affirms this strong link between quality and meaning when, for example, he writes:

> We follow, with apparently complete understanding, a tale in which a certain quality or character is ascribed to a certain man. But something said causes us to interject, 'Oh, you are speaking of Thomas Jones, I supposed you meant John Jones.' Every detail related, every distinction set forth remains just what it was before. Yet the significance, the color and weight, of every detail is altered. For the quality that runs through them all, that gives meaning to each and binds them together, is transformed. (196)

The change in the meaning of the narrative that accompanies the change in referent is grounded here in a transformation of qualitative experience. For Dewey, this "pervasive and underlying quality" (197) is grasped in the form of a "feeling, or impression, or 'hunch,' that things are thus and so" that indicates "the presence of a dominating quality in a situation as a whole."

Such impressions are not just, say, one way of enabling "us to keep thinking about one problem without our having to constantly stop to ask ourselves what it is after all that we are thinking about," but are necessary for understanding:

> Confusion and incoherence are always marks of lack of control by a single pervasive quality. The latter alone enables a person to keep track of what he is doing, saying, hearing, reading, in whatever explicitly appears. The underlying unity of qualitativeness regulates pertinence or relevancy and force of every distinction and relation; it guides selection and rejection and the manner of utilization of all explicit terms. (198)

A phrase like "the latter alone" suggests that qualitative experience structures and defines the meaning of sensations and concepts alike throughout the whole process of deliberation. Even so, Dewey affirms that inquiry involves the development of this hunch: "reflection about affairs of life and mind consists in an ideational and conceptual transformation of what begins as an intuition" (198-99). To say that a hunch or feeling is always operative in the background to structure all experienced and linguistic meaning is a strong claim, but it is fundamentally an explanatory claim, not a "myth of the given" style reduction of concepts to intuitions or appeal to intuition as justificatory. It would be

126

inconsistent with the principle of continuity and undermine Dewey's model of inquiry to claim that qualitative experience cannot be conceptualized. His point is rather that, just as objects are not perceived or understood in isolation from their context, so are our relations with others and ethical or aesthetic experiences defined by a shared situation.

But Dewey's approach to qualitative experience becomes a point of controversy between classical and neopragmatists. Gouinlock (1972) claims, I think rightly, that Dewey's naturalistic metaphysics is centered on his theory of quality, intended to "mark the general, pervasive qualities which distinguish one kind of situation from another" (7). Richard Bernstein (1961) critiques this foundation as incoherent, arguing that although qualities are supposed to be at once in experience and of nature, Dewey ends up applying the concept inconsistently:

> The distinguishing feature of experienced quality is that it can be mediated and funded. But when Dewey switches to discuss quality as an intrinsic possession of existences which do not necessarily enter into experiential transactions, he insists that qualities are unconditioned, they are precisely what is unmediated … Despite Dewey's claim for the continuity of experience and nature, he talks two different languages when he speaks of qualities as they function in experiential transactions, and qualities as the possession of natural existences independent of experience. (13)

Thomas Alexander (1987) replies that "by 'immediate' Dewey does not mean 'unmediated'," and that to say a quality in nature is immediate and unmediated is to refer "to the quality of the situation as a whole, inclusive of its determinate and indeterminate, cognitive and non-cognitive aspects" (80). To Bernstein's recommendation of setting aside Dewey's metaphysics in favor of his political philosophy, Alexander shows how the latter draws on the former in Dewey's philosophy insofar as the concept of the situation is dependent on that of quality. While this is correct, Bernstein may also be right to see in Dewey a lingering ambivalence between idealism and realism in how to

understand, for example, the externality of nature as an object of scientific inquiry. Regardless, I think there is a much more serious problem for Dewey's understanding of qualitative experience that has been less discussed by pragmatists: how a situation can be shared.

Although Dewey does not refer to his concept of the situation in his political writings, it is implicit in his concept of deliberation, which is politically significant. In *The Public and its Problems* (2016), Dewey says that "the improvement of the methods and conditions of debate, discussion, and persuasion" is "*the* problem of the public" (225), continuing:

> Until secrecy, prejudice, bias, misrepresentation, and propaganda as well as sheer ignorance are replaced by inquiry and publicity, we have no way of telling how apt for judgment of social policies the existing intelligence of the masses may be. It would certainly go much further than at present. (226)

His primary concern here is to refute an elitist or technocratic critique of democracy according to which the masses are insufficiently rational to be allowed to govern. In his introduction to the 2016 edition, Melvin Rogers expands on Dewey's concerns in a manner that shows their relevance to a contemporary context in which, again, democracy has come under attack:

> The integrity of democracy hinges on the extent to which the minority never feels permanently alienated from the process of decision making. Because the status of the minority is not perpetual, and as a result the minority does not exist under the weight of a tyrannical majority, the idea of political loss becomes an institutionalized reciprocal practice of decision making. This reciprocal practice is the deliberative "process" … Part of its function is to encode both the habits of reciprocity and mutual trust among citizens and between citizens and their representatives. To cultivate such habits is part of the process of mitigating the remainders of disappointment. (17)

The themes of continuity and the primacy of habit guide Dewey's defense of democracy, one he finds urgent precisely because of his rejection of a classical liberal view of

rationality as innate and independent of social factors. If one were pessimistic about the possibility of improving education or information-access, then Dewey's theory of habit might be deployed against democracy and so against Dewey. One might argue instead for rule by virtuous philosopher-kings who, say, act to shape the intransigent habits of others unable to transform themselves. Only the participation of individuals in local communities of deliberation can foster, against this hierarchical view of society, the sense that experts and masses are equals in a greater community.

But if deliberation is a response to a problematic situation, and to be in a situation is to have a specific qualitative experience, then it follows that for one individual to be in the same situation as another, both must share the same qualitative experience. Gouinlock (1972) touches on the problem of political conflict in his critique of Dewey's views on justice:

> Justice is the condition of the general social welfare; but there are many ways of construing the idea of social welfare. The notion of the *general* welfare admits of numerous possible interpretations and schemes of distribution. Few, if any, conditions satisfy the entire society; and different groups place differing values on existing laws, institutions, and practices: The common good is simply a will-o'-the-wisp. Hence in concrete cases some groups gain and others lose; and the disputing parties cannot agree on what would constitute an equitable agreement. These are facts of daily experience, and just how Dewey's notions of equality and justice could be applied to them is not clear.

Gouinlock concludes that Dewey "left much to be done on these topics" but that "constructive work on these problems must proceed in any case on naturalistic assumptions at least very similar to those of Dewey" (331). However, Gouinlock earlier affirms that political opponents share a situation. Citing the example of Black Americans in the civil rights movement, he writes:

> deliberation … might not take into account the responses of various white blocs. But – obviously – the whites are part of the problem. It is clear that to estimate

129

white response is not to venture beyond the situation, but to consider one of its most crucial components. Awareness of what is relevant to the situation is obviously vital to its successful transformation. (315)

This is the heart of the problem. Surely Gouinlock does not think that white opponents of civil rights share the same qualitative experience of the situation as its proponents. But it would be strange to say that they are not part of the "problematic situation" that civil rights protests, for example, aim to resolve. Gouinlock, in a manner consistent with Dewey, uses the terms "situation" and "context" interchangeably, but context is not necessarily qualitative. The context of the civil rights struggle is historical. That context does manifest in social and material particulars, awareness of which might trigger hunches or impressions of racial injustice that lead to further inquiry. But, as perhaps in the example of a white moderate, this awareness could instead be book-knowledge that first came as a complete surprise and then is consciously and reflectively applied to interpret social conflict despite the absence of hunches or impressions of racial injustice. Not all background context is grasped through background cognitive processes.

Consider, as an example of a qualitatively ambiguous situation, a possibly offensive joke. Imagine a gathering of three persons in which A tells a joke that is arguably at C's expense, B intuitively takes the joke to be innocuous and laughs, and C is offended and confronts A. A may feel that C's reaction is the problem. C may feel that A's action is the problem. B may feel that the problem is figuring out if C's reaction or A's action is the problem. First, it is plausible that there is no shared qualitative experience in this situation. But second and more importantly, the problem may precisely be *the absence of a pervasive unifying quality*. The disagreement arises because everyone's initial felt intuition or hunch is divergent unto incompatibility. If a

problematic situation is defined in terms of the presence of a pervasive unifying quality, then the idea that a problematic situation can be about the absence of such a quality is unintelligible.

Now, one might respond that there would be a pervasive unifying quality of, say, tension. First, I suspect it would be implausible to specify a *uniquely identifying* pervasive unifying quality when each party's felt problem is different to, as for A and C, the point of incompatibility. But, second, this presumes good faith. Let us assume that A is of a higher social status or wields more social power than C. Now, imagine "A_2," who acts in bad faith. A_2 is aware that C's social position is precarious, perhaps because C is a member of a stigmatized group. A_2 deliberately mocks C in ways that will seem to B to be innocuous, provoking C into reacting explosively, which will seem to B to be unreasonable. B may then come to see C as untrustworthy and A_2 as more trustworthy. If C does not react, then their failure to respond to A_2's mockery may entrench the sense of their lesser status in B's mind. If A_2 aims to entrench their higher status in B's mind, whether C fights back or not, this is a win-win scenario. In this case, A_2 might plausibly not be feeling any tension at all. This is a calculated and deliberate action meant to seem casual and accidental, all to weaken C's case if they respond.

One might then respond that the appropriate move for B to is take C's side. But this is only the case if A actually is of a higher social status and is punching down. C may just think they are of a lower social status and react as such but be mistaken. According to a 2016 poll, around 45% of self-identified conservatives, twice as many as self-identified liberals, claim that men are persecuted just for being men. Imagine that "A_3" is a woman who has made an offhanded joke about men and C is a man who wrongly believes that

society is more sexist against men than against women. In this case, A_3 is punching up the social status ladder, but C is reacting as though A_3 is punching down. This might be sincere, or it might be a tactical defense of social power that might otherwise erode. Either way, even if A_3's joke was wrong, it might be better for B to be circumspect or to try to talk C down than to side with C against A.

Here is the kicker: A_3 might also be acting in bad faith, but for the benefit of the group. C might be a puffed-up incompetent man who bullies others while seeing himself as the victim. A_3 might be deliberately trying to make C seem ridiculous in front of B for basically the same reason as A_2: to make C seem less trustworthy and A seem more trustworthy to B. A_3 correctly assesses that C's authority is a malign influence on the group that, if countered, would benefit B and others. But if C is truly undermining the group yet too socially powerful to challenge directly, then A_3's trolling might be the only effective method of countering C's influence. C would be akin to a tyrant lashing out at the jester who, in revealing his absurdities to the people, lessens his authority.

Finally, imagine oneself as B. How does B know if A is A_2 or A_3? What possible information within the situation itself would be a trustworthy guide? Does C's anger mean that C is oppressed or deluded? Do A's protests mean that A is innocent or a cunning manipulator? If A or C's self-description is a charade, how confident is B that they can see through it, ever?

This problem is what I call a "qualitatively opaque problem," the correlate of a qualitatively ambiguous situation, and the answer to it is that B must quickly turn away from intuition and toward reflective knowledge of context. C's anger might be very convincing. But if C is a white man launching into a tirade about how "not all men" are

oppressors, it is B's knowledge of historical and cultural context that will frame this anger as probably reactionary rather than just. C might still be perfectly sincere. But this information is vital for B in discerning to what C's sincere anger responds: an erosion of a traditional but unjust hierarchy rather than, say, a natural and just hierarchy. If C can be reasoned with in this situation, it may be by first distancing him from his intuitive or emotional reaction rather than by attending to what it seems to reveal, which is false. In a qualitatively opaque problem, even beginning with one's intuition or hunch may be likely to lead inquiry astray rather than to its eventual refinement through deliberation. Such a strong and misguided reaction is better nipped in the bud before it blooms into confusion. The response to a qualitatively opaque problem is thus to put reflection before intuition and to turn to a more idealized, less hunch-guided form of normative discourse.

Pleasant birdsong is a fierce competition. Tranquil trees fight for the sun. These facts prove Dewey's point about inquiry. If aesthetic pleasure in birdsong motivates an inquirer, who assumes birdsong to be essentially as mild or peaceful as it seems, to a scientific conclusion that reveals that they began in error, so much the better for aesthetic pleasure as a source of knowledge. I affirm this point. I do not think that, say, birdsong having the instrumental value of motivating knowledge-seeking somehow tarnishes the pleasure in it that is experienced as having intrinsic value. I am inclined toward the pragmatic view that more value is better. I even agree that most inquiry, generally, may begin successfully with hunches and intuitions and that reflection may safely follow. I thus see my modification to Dewey's epistemology as minor.

However, I also take my contribution to scholarship on pragmatism to be major. The concepts of the qualitatively ambiguous situation, in contrast to a situation defined

by a pervasive unifying quality, and its correlate, the qualitatively opaque problem, represent a distillation of the array of critiques aimed at classical pragmatism regarding its views of normativity, social power, and tragedy. To think that every problematic situation is defined by a pervasive unifying quality is to express excessive optimism unto naivety about the ways in which values may (or may not) become norms, social power may oppress, and tragedy may arise from habits so intransigent as to seem like fate.[32]

I do not think that Dewey is naïve. His educational and political projects are much too grounded in realism and concreteness. But I do think, as I will discuss when I return to Dewey in Chapter 5, that philosophers who follow Dewey, like myself, risk naivety without a concept of a qualitatively ambiguous situation that calls for reflection before intuition and a concept of a qualitatively opaque problem that calls for a more idealized normative discourse to override from the outset a non-idealized chorus of hunches. That said, I note that this may remain a fundamentally Deweyan account if the imperative for B becomes to *form a habit* of beginning deliberation in response to qualitatively opaque problems with reflection rather than intuition.

---

[32] I hope to later engage more fully with Dewey's theory of power (see, e.g., Wolfe 2012) and Cornel West's critique of the pragmatic view of tragedy (see, e.g., Ruprecht 2017). I strongly suspect that the lack of a theoretical account of qualitatively opaque problems motivates the assessment that Deweyan pragmatism does not acknowledge or counter the insidious ways in which manipulative uses of power can lead to tragic backsliding, as when the United States followed post-Civil War Reconstruction with the white supremacist "Redemption" and the rise of Jim Crow. I think that there is an analogy to be made between B and the "white moderate" of Martin Luther King Jr.'s "Letter from Birmingham Jail" (1963). A version of B who sides with A2 and ignores the historical context of social power by going along with the existing hierarchy may, annoyed, similarly wish that C would just quiet down. In the context of an ethics for the depressed, as I will discuss in Chapter 5: B is 1) failing to value the higher (or less partial) values of compassion and fairness over the lower (or more partial) values of loyalty and deference to authority and 2) lacking the historical knowledge required to accurately evaluate qualitatively ambiguous situations, which B would be motivated to acquire by those same higher values of compassion and fairness. If tragedy implies a lack of responsibility, then this B is more contemptible than tragic. But if B values higher values and yet is duped into opposing higher values by mechanisms of social power in a context of qualitative ambiguity, this may be tragic.

Before I turn back to Ratcliffe's phenomenology to show how my defining

Dewey's concept of habit as representing a pre-intentional state reveals the limitations of

Ratcliffe's anti-conceptual definition of existential feeling and his refusal to distinguish a

correlate concept of existential dispositions, I pause to consider two questions. First, why

does this conflation of experienced and historical context into a single concept of the

situation occur in Deweyan pragmatism? Second, why are neopragmatists attentive to

Dewey's theory of quality like Bernstein focused on a critique of Dewey's naturalistic

metaphysics instead of the normative tensions between Dewey's epistemology and

politics arising from how the concept of the situation is defined?

A defender of Dewey could object to my interpretation by emphasizing, as I have

noted, that Dewey does not explicitly state that political adversaries share a problematic

situation, in which case this issue has been little discussed because I am misguided. But I

think that Dewey's concepts of deliberation and inquiry, which are continuous (of course)

between his epistemology and his politics, are too closely tied to the notion of a response

to a "problematic situation" for this response to amount to much of a defense. That said,

introducing distinctions between different aspects or kinds of situation or context, some

qualitatively unified and some qualitatively ambiguous, preserves the main components

of Dewey's epistemology and politics. Here, I want to highlight the most important

takeaway for this project from this story of scholarship on Deweyan pragmatism: the

conflict between classical and neopragmatists is fundamentally over how philosophers

should discuss normativity. It is understandable that professional philosophy should go

the way of the neopragmatists in strictly distinguishing normative argument from

descriptive accounts, making argument the business of philosophers and viewing

something like an "exhortative metaphysics" with suspicion. But the Deweyan approach in a broad sense resonates with the needs and goals of an ethics of the depressed.

Dewey's defense of egalitarian democracy against hierarchical elitism or technocratic politics is guided by a vision of gradual progress in which all that which is treated as vulgar or parasitic, whether categories or persons, is a constructive contributor. For the liberation of self and society alike, habit is as instrumental as thought, the body as the mind, the masses as experts, crafts as arts, and on and on. He is right to see how an epistemology focused on education, sensitive to its transformative power and to the challenges posed by entrenched habits and the detachment of elite policymakers, entwines with an egalitarian politics. And, as Richard Rorty (2004) writes, Dewey is justly inspired by Darwin's naturalism to "see cultural evolution as on a par with biological evolution – as equally capable of creating something radically new and better" (23). Because of his focus on rehabilitating phenomena like the "hunch" as substantive contributors to processes of inquiry, against the presumption that a gut instinct has little epistemic value, it would make sense that he is less attentive to cases of inquiry not initiated by hunches or not guided by a shared sense of a problem.

The linguistic turn's emphasis on the difference between the normative and the descriptive, between giving reasons and accounting for causes or weaving a narrative, is not just correct but important. Defining the paradigm case of philosophy as the good faith exchange of reasons and not as, say, offering descriptive accounts with normative implications, may help in, for example, resisting manipulative attempts to shape the views of others without their awareness or consent. The way this emphasis constrains professional philosophy may be the product of ceding subject matters of inquiry to other

disciplines in the sciences, but it also represents noble ideals of rigor and humility. Still, as Bernstein (2010) later argues, this "severely limits the range of human experience (historical, religious, moral, political, and aesthetic) that should be central to philosophical reflection" (152). Much that is under the rubric of "causes," like compulsions resistant to argument, becomes on this view necessarily unphilosophical: if it is not a matter of "reasons," it is a matter for scientists instead. The distinction between causes and reasons or the descriptive and the normative may be maintained without surrendering so much on the question of what counts as philosophy. Perhaps if descriptive accounts with normative implications, or at least those that are clear about what they are, were better tolerated in professional philosophy, the depressed would find it a more welcoming discipline.

I suspect that Dewey's naturalistic metaphysics became the focus of debate in the pragmatist tradition, rather than this internal tension between his epistemology of situations and his politics of deliberation, because thinkers like Rorty and the early Bernstein objected more broadly to the relevance of metaphysics to normativity and the accompanying risk of conflating causes and reasons. The rejection of appeals to nature as justificatory of norms, common to Anglo-American analytic philosophy and European post-structuralism alike, understandably clashes with the ambiguous status of Dewey's metaphysics as both a straightforward descriptive account of the features of experience and, also, a sort of after-the-fact narrative of how to think about the world consistent with his epistemology and politics. The Rortyian position is to, while appreciating humanistic storytelling, resist any semblance of grounding politics in metaphysics. Dewey was, I think, less concerned with the appearance of fallacious justification and more with

137

countering anti-democratic narratives with narratives of his own. And indeed, the Deweyan naturalist point that habits of good empirical inquiry may also be habits of good normative inquiry is very different from the claims of figures like Steven Pinker, in whom Rorty (2004) diagnoses "philosophy-envy" (22), that better normative conclusions will follow somehow from better scientific conclusions. Today, the rise of moral psychology suggests a renewed partnership between philosophy and science rather than just a division of labor, though as my account of its treatment of depression implies, the philosophical side of this partnership seems to remain hung up on matters that have traditionally concerned the analysis of language, meaning intentional states that may take the form of propositions. This history offers clues to a model of philosophy in which an ethics of the depressed, responding philosophically to a dilemma resistant to reason-giving or argument, is both successful and rigorous. It offers a thoughtful broadening of what counts as properly philosophical, material circumstances like departmental budgets permitting.

A Deweyan understanding of habit coupled with a modified concept of the situation that is only contingently emergent from qualitative experience suits the habit example of the pre-intentional. A habit is more than a reflexive response or pattern of action: it is a disposition that at once constrains or focuses attention and provides a cognitive shortcut to a related pattern of action. To say that it provides a cognitive shortcut is to say that it is able and tends to bypass reflection, increasing efficiency at the possible cost of overlooking new or anomalous information. Its bypassing of reflection means that one's awareness of a situation might only reach the level of a hunch or feeling before a habit kicks in and one begins acting in response, though a habit may also refer to

138

any pattern of mental association whether or not there are any experienced cues involved. My example of a habit of feeling suits this Deweyan understanding of habit as a mediator between intuition and action: a depressed or anxious person might feel a strong sense of being threatened and then, seemingly without having the chance to decide on the reliability of this intuition, interrupt all other action or thought to scan for threats in a defensive posture. Understood this way, a habit is a good candidate for a pre-intentional state: it does not aim toward an object, but it actively assists other mental states in their aiming toward their objects. Still, one might reasonably reject this framing and insist that mental states are either one or the other and, even in the complex case of a Deweyan understanding of habit, such states are fundamentally nonintentional. I will return to this objection in the conclusion of this chapter when I argue for a modification to Ratcliffe's definition of the pre-intentional.

But one more reason why the full account of Dewey's theory of habit is relevant to a discussion of Ratcliffe's theory of existential feelings is that Ratcliffe makes a similar misstep: his descriptive account is led astray, overstating the influence and independence of existential feelings, seemingly by a normative concern to center the experiences of the depressed by countering a philosophical narrative in which the pre-intentional is reducible to intentional states. As I have argued in line with Ratcliffe in the previous chapter, reducing questions about judgment or motivation to the terms of belief-desire psychology neglects the possible structural effects on both by states like moods or habits, thereby leaving depressed persons with few resources for understanding the effect of depression on either. I have proposed a quasi-belief model of demoralization that is consistent with belief-desire psychology, but pointed out that there is an equally plausible

model for demoralization, the habit model, that does not begin with a change in beliefs or desires. The habit model plausibly fits into Ratcliffe's understanding of the pre-intentional, the paradigm case of which for him is existential feelings. But, as I will now show, the epistemic function of existential feelings has come under serious criticism and his response has been inadequate. By introducing a distinction that Ratcliffe rejects between existential feelings and existential dispositions, I intend to clarify the bounds of the pre-intentional and render the concept both more compatible with belief-desire psychology and more practically applicable for depressed persons making sense of their experiences.

3. Existential Feelings and Existential Dispositions

It is harder to make the case for the pre-intentional as a category of analysis, and thus to discuss existential feelings in the context of analytic moral psychology, when Ratcliffe insists that existential feelings *are* pre-intentional structures and not, say, ways of becoming aware of pre-intentional structures. Ratcliffe explicitly rejects a distinction between existential feelings, which are necessarily felt whenever they are present or operative, and what I call existential dispositions, which would be pre-intentional structures that might be apprehended by existential feelings or might not be. Ratcliffe is clear that for him, wherever there is a pre-intentional structure, there is a feeling. Jussi Saarinen distills the challenges presented by Ratcliffe's approach by identifying three theoretical commitments made by Ratcliffe that are vulnerable to criticism: the dependency, bi-directionality, and non-conceptuality theses.

The dependency thesis is that, according to Ratcliffe, intentional states always depend on a pre-intentional framing of possible objects to which the states may be directed in the form of an existential feeling. The problem is that feelings are felt: they must be at least minimally present in our awareness in order to be feelings. Saarinen raises the example of the emotion of fear being dependent on "the possibility of being threatened" and asks if this possibility makes sense as a "felt bodily sense." Ratcliffe presents such feelings as "ordinarily inconspicuous" but, as Saarinen points out, they can never be "completely inconspicuous": "we may ask whether it is plausible to designate as 'feelings' aspects of the experiential background that rarely, if ever, shift into the field of reflective consciousness" (7). This question also prods at the third existential category Ratcliffe has layered onto the noetic-noematic or *Leib-Korpor* distinction. It seems to me that Ratcliffe has taken a second transcendental step: *Leib* structures the way that *Korpor* is encountered and existential feelings structure the way that *Leib* is encountered. Setting aside the risk of taxonomical redundance here, just how many feelings are we supposed to be having at any given moment? Bodily feelings associated with anger may influence the perception of the object of anger, for example, but the "possibility that objects in general can be disclosed as offensive or provocative" (11) is one possibility among many. To suggest that this possibility is necessarily felt, however minimally, implies a very crowded emotional life.

The bi-directionality thesis concerns if intentional states can influence existential feelings: Ratcliffe seems committed to this being true, but it is not clear how to make sense of that possibility. Considering the question of how a self-narrative or an interpersonal interaction influences one's existential feelings, Ratcliffe asks, "how could

an intentional state somehow 'act upon' its own conditions of intelligibility?" and answers "it simply happens, just as existential changes can happen when one is sick, tired, or intoxicated" (2015, 151). I would agree that existential changes often do seem to "simply happen." But Saarinen points out that what makes this question difficult to answer is an ambiguity in the concept of existential feeling: "much hinges on whether existential feelings are identified as background affective styles or consciously occurring episodes" (2018, 8). How intentional states might affect their own conditions of intelligibility will be understood differently if those conditions are, to pick from previous examples, a more concrete feeling of existential hope or anxiety or a more abstract sense of the possibility of finding objects in general offensive or provocative. My anxiety might foreclose my capacity to encounter a prospect for action as possible for me, but my belief that my anxiety is a misleading compulsion might facilitate at least some action – even if indirect, like deep breathing – that eases the anxiety. But if the pre-intentional framing is "minimally felt," it would be difficult to take it in the moment as an object of an intentional state like a belief, so bidirectionality would presumably be understood differently. This confusion arises because existential feelings are presented both as constant backgrounds and vivid occurrences.

The last challenge arises from Ratcliffe's claim that existential feelings are non-conceptual: judgments are intentional states and existential feelings are pre-intentional, thus their sense ostensibly does not arise from the application of a concept but is directly felt. For Saarinen, this claim is most dubious in the context of existential guilt, which of all the existential feelings is the most straightforwardly evaluative: it arguably implies "at least a rudimentary grasp of concepts pertaining to moral agency, e.g. right and wrong,

reproachability, blameworthiness, normative ideals, and so on." Even granted that there might be such a thing as a "pure" bodily feeling like "feeling alive, fresh, or tired," the conceptual complexity of intentional guilt suggests that existential guilt, and perhaps other existential feelings, "might require certain conceptual/evaluative schemes as conditions for their own intelligibility" (10). I would go further than Saarinen here and raise the concern that Ratcliffe risks in this "non-conceptual" claim an intuitionism vulnerable to a Sellarsian critique of the myth of the given: in brief, that correctly knowing or identifying what experience is "guilt" depends on correctly applying the concept of "guilt," information which no intuition or feeling intrinsically contains. Surely there is at least some conceptual mediation involved in the experience of an existential feeling.

I sympathize with what I take to be a motive of Ratcliffe's for these positions, that is, wanting to avoid reducing existential feeling to a kind of epiphenomenon of the concepts or dispositions that "really" define the experiences of depression. I think it is true that there is a category of existential feelings that are underexamined because philosophical psychology centrally concerns intentional states – beliefs, desires, action-preparatory emotions – and object- or goal-directed judgments and actions. These theoretical priorities are complicit in the misunderstanding of experiences of depression. I also think, as stated before, that existential feelings could have some direct effect on one's sense of what is possible in, e.g., taking up "mental bandwidth" as cognition-interrupting vivid occurrence. But insisting categorically on the priority of existential feeling to seemingly most or all judgment will fail to ensure its relevance to moral psychology and only render its definition confusing and its application dubious. A

plausible alternative is that existential feelings generally do not themselves define our experiences of possibility but instead make us aware of dispositions that otherwise operate beneath our awareness to define our experiences of possibility. The burden is on Ratcliffe to show why this interpretation of existential feelings should be rejected.

Ratcliffe's brief response to Saarinen's criticisms is inadequate and inconsistent. Ratcliffe reaffirms his rejection of accounting for pre-intentional framing of possibilities in terms of dispositions: "differences in the types of possibility to which one is receptive … are integral to experience rather than consisting of non-experiential dispositions" (2020, 252). He defends an expansive definition of existential feeling that includes both what Saarinen calls "background affective styles" and "consciously occurring episodes" by offering the example of a self-doubting teacher experiencing a "pervasive sense of unease" when they enter the classroom. "This style of anticipation has a degree of consistency and coherence that warrants our calling it a distinctive kind of 'feeling,'" Ratcliffe continues, but because this style manifests so broadly "in one's experiences, thoughts, and activities (or lack of activities)," it is not "a singular psychological entity" and "there is no clear line to be drawn between accessing the feeling itself, recognizing it through its pervasive effects, and somehow inferring its presence" (258). In response to the challenge that, if existential feelings are the precondition of possible experiences, there would have to be innumerable existential feelings, he distinguishes a "dynamic experience of the possible *allow*[*ing*] for certain kinds of emotional response and not others" from "the stronger claim that a sense of *what, exactly, it allows for* must be a constant ingredient of existential feeling." In other words, "one can be open to

144

possibilities of type p without having an additional, continuous experience of being open to possibilities of type p" (257).

But Ratcliffe's commitment to constraining "differences in the types of possibility to which one is receptive" to the realm of feeling undermines his own points. First, style is more than a feeling. One's "style of anticipation" could have "consistency and coherence" without being identified with a feeling: it could operate as a set of belief-like assumptions or habitual dispositions. Even if we perceive style in ourselves and others in a "pre-reflective" manner, as Merleau-Ponty argues in his *Phenomenology of Perception*, one might not "feel" one's style of anticipation but infer at least some aspects of it from one's own patterns of perception or action. Second, to say that style is not a "singular psychological entity" seems to endorse the notion that it is not just a feeling, in contradiction with the position that pre-intentional structures are existential feelings and not dispositions. To say that "no clear line can be drawn" between, e.g., feeling a feeling and recognizing or inferring what it is smuggles in intuitionism without openly endorsing it: naming a feeling means employing a concept. Perhaps the person experiencing an existential feeling might struggle to draw a line between feeling and knowing, but a theoretician can and must. Finally, if we are distinguishing the content of existential feelings from their consequences for our openness to possibilities, those consequences become both a separable and more relevant factor in our analysis. Grant that a feeling of anxiety does not involve some feeling of the content of what possibilities it will allow or disallow. If those possibilities constitute the pre-intentional, then the pre-intentional is something other than a feeling.

Ratcliffe's conflation of feeling with what is felt is an example of what I call a fallacy of the transparent body. For phenomenology since Husserl, the body or embodiment operates as a transcendental schema that is apprehended rather than deduced, though how this apprehension is possible – if, for example, it requires a complete suspension of the "natural attitude" that typically structures perception – is widely debated. In theory, this body schema is always operative but only potentially apprehended through phenomenological attention or analysis. In practice, however, phenomenologists sometimes treat the "bodily" status of some aspect of the body schema as a justification to treat the schema as *itself constantly perceived*, as though phenomenology is not required to uncover it. This conflates the form and the content of perception: a mode of perception, like bodily feeling, becomes part of what is perceived, like the "mood" or "atmosphere" as grasped through or interpreted by one's dispositions. It thus also conflates the mediate and the immediate, which is to say, the background and the foreground: what structures or mediates the "given" in experience is also treated as "given," always already present to the subject. Ratcliffe's identification of existential feelings with dispositions is an example of this fallacy. There is no good reason to think that every transcendental structure shaping thought and perception is simultaneously felt, however minimally that feeling is supposed to be present, let alone that such feelings may be identified without the application of concepts. To insist otherwise is to insist that what is significant must also always be present.

The fallacy of the transparent body, then, is an example of the "privileging of presence" which deconstructionists like Jacques Derrida critique in phenomenology. The irony is that phenomenology has since its inception been attentive to its own ambivalence

regarding presence in the Husserlian example of the "touching touch," the experience of

two of one's own fingers touching one another, in which one perceives oneself as both

touching and being touched. In his final and posthumously published work, *The Visible*

and the Invisible (1968), Maurice Merleau-Ponty refers to the touching touch as "*a*

reversibility always immanent and never realized in fact" (148). What is at stake in how

phenomenologists understand the touching touch becomes clearer when one considers

that the distinction between foreground and background and the claim that the

background structures the foreground is, as in Dewey's pragmatism, a core insight of

phenomenology. Indeed, what justifies the need for phenomenology as a method is that

the background is in the background and thus may escape attention. The touching touch

is an event in which the background becomes foreground, becomes present, and so it can

serve both as an exciting example of the importance of phenomenology and a threat to its

own necessity. Jacques Derrida (2005), in his reflections on Merleau-Ponty's discussion

of the reversibility of the touching touch in *On Touching—Jean-Luc Nancy*, expresses

"regret that he [Merleau-Ponty] was unable to carry out a more powerful reformalization

of his discourse" and claims that despite the ambivalence of the touching touch, he was

"always, *in fact*, and all things considered, *preferring* 'coincidence' (of coincidence with

noncoincidence) to 'noncoincidence' (of coincidence with noncoincidence)" (211). Even

when phenomenology is dedicated to the uncovering of structures that are not present,

there remains an enduring temptation, Derrida thinks, to reduce all objects of

phenomenological investigation to the "given" with which phenomenology begins. This

is what he calls the "haptocentrism" of a desire for absolute self-contact that denies that

what structures our perception or thinking may be inaccessible to us except indirectly, via traces left behind.[33]

I can resolve Saarinen's challenges by rejecting Ratcliffe's fallacious conflation of existential feelings and dispositions and instead understanding existential feelings as pointing the way toward the pre-intentional rather than as identified with it. If the pre-intentional is understood as habits, dispositions, or quasi-beliefs that are not necessarily even minimally felt, then intentional states need not be understood as necessarily somehow dependent on feelings. The bidirectionality of existential and intentional states could be analogous to that of habit and action, the former having an initial impulsive and preparatory shaping effect upon the latter but being nevertheless malleable and potentially receptive to deliberate (re)formation. Similarly, existential feelings could be non-conceptual in the sense that, like habits, they may have consequences for what actions are likely to happen without requiring the mental representation of an object or goal, while still implying some sort of conceptual apparatus as part of their own conditions of intelligibility as Saarinen says. This is a weaker concept of existential feeling that may not seem to speak to the strength with which they are experienced. But how something feels is not necessarily how it is best described, and indeed, part of the point of applying the concept of existential feelings to depression is that they can mislead us, e.g., in our sense of what is possible for us. And if existential feelings remain our best

---

[33] If we take Derrida's critique seriously, then perhaps this haptocentrism is itself the trace left behind of Husserl's claim to apodictic certainty for phenomenology, as being not just a refined mode of introspection but a rigorous science of consciousness capable of grounding empirical sciences via an account of the constitution of the world by the subject. Even as Husserl's successors like Heidegger and Merleau-Ponty, and eventually Ratcliffe, set aside the ambition for certainty and the question of phenomenological method, the inclination to make background structures "present," solid and indubitable, remains.

way of becoming aware of at least some of our present pre-intentional dispositions, they remain extremely important.

Ratcliffe's account of existential feeling makes a strong case for distinguishing the pre-intentional from a cluster of intentional states. Existential hope and guilt are vivid examples of the experiential unity and broad consequences of what I have called the shroud effect of depression. As I have said, "pre-intentional framing" effectively refers to a category of mental states that modify other intentional states and could be understood either as 1) intentional states directed toward other intentional states or 2) non-intentional habits or dispositions that happen to interfere with intentional states. Crucially, my redefinition of existential feeling as a way of becoming aware of the pre-intentional, and not the pre-intentional itself, makes it compatible with my second-order quasi-belief account of depressive loss of motivation: existential feelings, in making us aware of what I have broadly called existential dispositions, could make us aware of operative quasi-beliefs. This interpretation of the pre-intentional, by defining it in terms of belief, makes it less disruptive to the belief-desire model of psychology.

However, two implications follow. First, as concluded in the previous chapter, if we lack the epistemic access required to tell the difference between a second-order quasi-belief and a non-intentional habit, then the commitment to understanding motivation in terms of intentional states seems in need of further interrogation or justification. Second, given the pervasive effects of existential feelings, such as on experiences of time and interpersonal relations, the concept of quasi-beliefs might be more broadly applicable in making sense of, say, entrenched or recalcitrant habits, specifically understanding how

149

they might have a kind of propositional content or operate as though they do without being reflectively adopted or endorsed.

Ultimately, the problem with Ratcliffe's account for an ethics for the depressed is that his refusal to accept that existential feelings have conceptual content means that he cannot give a prescription for existential hope. When he reflects on how intentional hope might generate existential hope despite the former having conceptual content and the latter being a transcendental condition of the former which somehow does not, he can only say, as quoted, "it simply happens" (2015, 151). My point is not to belabor my previous critiques. As I have said, I am sympathetic to where I think Ratcliffe is coming from in wanting to position pre-intentional states as more than epiphenomena of intentional states. In my redefining the category of the pre-intentional as a cluster concept of "existential dispositions" that includes intentional states like quasi-beliefs and intentionally ambiguous states like habits of feeling, I paint a similar picture. For example, quasi-beliefs may have a practical priority over beliefs in conflicting with them but being more resistant to refutation. My point is that Ratcliffe's erroneous strategy for establishing the priority of the pre-intentional leaves it mysterious as to how a depressed person might gain existential hope. I, however, may clarify: the answer lies in indirect responses to quasi-beliefs, like mindfulness or gentle questioning, and counter-habits developed in response to habits. In both a literal and metaphorical sense, a metaethics must give hope for an ethics to get a grip.

4. The Pre-Intentional as a Condition of Motivation

Depression is commonly characterized by a loss of interest, which has made it relevant as an example of loss of desire to the debate over moral judgment internalism and externalism in metaethics and moral psychology. As we have seen, this application of the concept of depression is confused on several levels. First, the suffering and alienation associated with depressive loss of interest are better understood in terms of a compulsive change in belief rather than loss of desire, even though the latter presents itself as an intuitive explanation. Seeing the symptom of loss of interest as emergent from a quasi-belief in something like the futility or threateningness of action can explain how the other most common symptom, guilt or feelings of worthlessness, can be entangled with that loss of interest: that futility or threat can, for example, be experienced as arising from one's perceived lower moral status. This account of demoralization is more accurate and complete. Second, depression does not seem to cause a loss of interest in "moral" or other-directed concerns at all, but rather in "prudential" or self-interested concerns. One might conclude that depression is thus irrelevant to debates in metaethics and moral psychology insofar it has nothing to do with moral judgment. I conclude instead that depression serves as an example of how reducing questions about motivation to questions about judgment, and component intentional states like beliefs and desires, is a mistake.

To moral psychologists: there are "pre-intentional" dispositional conditions of the possibility of motivation like habits and moods, undertheorized due to lack of attention to existential feelings in discussions of emotion, and whether they are understood as non-intentional or second-order intentional states, they are significant to how best to theorize motivation. This may be a way of rethinking the motivational role of virtues instead of

through the concept of "besires." And it may also strengthen what I take to be Swartzer's radical notion: that even desires should not be seen as intrinsically motivating states. But as to whether "desires intrinsically motivate but can get nipped in the bud by the pre-intentional" or "desires can't even get off the ground without certain pre-intentional dispositions" is the more accurate picture, I am not prepared to say.

To depressed persons: the practical problem of moral motivation in depression is not what it seems. It may appear that your goals are empty or impossible, that you are wretched and contemptible, and that these are somehow linked. This link is not what it might present itself to be: you are not choosing this state, you are not morally failing, and the perceived futility of your desires is not "caused" by your moral failure because it does not exist. And you do have desires: the simple fact that you hurt is evidence. What cuts these desires off from action are claims you find yourself making: that success is impossible, that failure will have terrible consequences, that this is true only for you and no one else. Even when you know these claims are false, that knowledge will not stop you from making them. If you cannot argue yourself into submission, do not despair. You are still reasonable and still have agency. Quasi-beliefs are ridiculous: there is nothing shameful in overcoming them indirectly, with patience and forgiveness toward yourself. And habits are intransigent even when not reinforced by an element of compulsion. So, the practical problem of moral motivation is not motivating yourself to be moral, even if it will seem that way to you at first: it is quieting your quasi-beliefs or interrupting your habits for long enough to experience yourself judging risks and rewards, bads and goods, with the equanimity you already possess.

An ethical theory constructed with depressed persons in mind is contending not with a morally intransigent agent but someone vulnerable to self-castigating patterns of thinking, who will flinch from self-interest, throw themselves into altruistic thinking or action, and despair all the while. Such a person may benefit from an explanatory context like the concept of the pre-intentional and its attendant examples of quasi-beliefs and active habits, which must be argued for, but not directly from arguments, against which quasi-beliefs and habits are recalcitrant. Thus, a depressed person's inclination to seek reasons to go on may be straightforwardly futile insofar as they can never been given good enough reasons to stop being depressed, but become indirectly useful in encountering a way of thinking relevant to their experiences of themselves as agents.

Having reframed the concept of depressive loss of motivation, in the next chapter, I move to the concept what depressed persons may perceive this motivation as directed toward: moral obligations. I argue that a moderate ethical theory, or what I think of as "ethics as a reliable guide," can foster depressed persons' sense of agency by not relying on the moral-prudential distinction that is operative in procedural ethical theories that assign moral obligations a special status. I review the literature on psychological realism, focusing on Bernard Williams's critique of "the morality system," Owen Flanagan's critique of Williams and alternative theory of minimal psychological realism, and David Wong's assessment and reformulation. By relying on Shaun Gallagher's account of the sense of agency in intention-formation, I argue that while depressed persons are motivationally demoralized, they are also at once deliberatively "hypermoralized," unable to maintain a distinction in deliberation between merely prudential interests and moral obligations. Insofar as formalist or procedural ethics involves a decision procedure

153

that requires agents make such a distinction so as to understand what obligations are moral and thus overriding, depressed persons, though otherwise rational, will be unable to conclude moral deliberation and will experience further, and intensified, devastation. The twin challenges of demoralization and hypermoralization for formalist and proceduralist ethics will bolster my argument that an ethics for the depressed is a nonformal value ethics.

CHAPTER IV: DEPRESSION, DELIBERATION, AND MORAL OBLIGATION

A strict ethical theory that proscribes making an exception for oneself in an egoistic sense, like by placing one's self-interest above one's ethical obligations, risks exacerbating the suffering associated with the depressive compulsion to make an exception for oneself in an anti-egoistic sense. Depressed persons, inverting the egoist, exaggerate the demandingness of their own ethical obligations, but not that of others, and so feel their own moral worth to be especially lowered. The risk of strictness is clarified by the preceding discussion of depression and motivation. Recall that a depressed person does not lack the desire to act altruistically – this may even be intensified – but experiences themselves as incapable of successfully acting, in whatever terms we understand that experience (as quasi-beliefs or habitual feelings). Such a person may, like the egoist or skeptic, seek a strong justification for their ethical obligations, but not because the depressed person will otherwise forswear these obligations. Rather, someone reduced to rumination or threat-prediction out of an felt incapacity to successfully act likely hopes on some level that a strong justification will bring with it some motivational force that will overpower the motivational block, one born of a compulsion resistant to argument but manifesting as a spiral of questioning that seems to call for argument. It is plausible that an ethical theory designed to be strict, or what Samuel Scheffler (1992) calls a "stringent" theory, in having a precise procedure or objective foundation or special category of ethical obligation, will call to a depressed person yearning for a leap from theoretical justification to practical motivation and, at the same time, feed the depressed person's self-castigating downward spiral. Such theories seem designed to counter the risk of agents not going far enough, not of going too far, if only in their heads.

155

This concern, that there exists a risk that the application of a strict ethical theory aggravates the effects of existential guilt, resembles the objection to modern morality that emerges from the theory of psychological realism, associated with Bernard Williams and Susan Wolf. These psychological realists argue that modern morality is excessively demanding and alienates agents from their personal projects and identities. This allegedly excessive demandingness has a cluster of causes in different conceptual mistakes about obligations, practical necessity, voluntariness or the will, and justice. In Williams's view, however, the central issue at stake is the rejection of moral luck: that is, the claim that a properly constructed morality does not evaluate the moral status of an agent or agent's actions based on circumstances beyond that agent's control. The meta-level unfairness of moral luck implies that a strict ethical theory ought to eliminate it. A moderate ethical theory, which does not aspire to offer a precise procedure or objective foundation for justifying claims, might instead endorse Williams's assertion in the last paragraph of *Ethics and the Limits of Philosophy* (2006): "the idea of a value that lies beyond all luck is an illusion" (196). The absence of an entirely fair procedure for evaluating anything, including ethical evaluations of acts or persons, would undermine any strict ethical theory. Moreover, according to the proponents of psychological realism, the effort to produce that fair procedure to determine what is valuable "beyond all luck," or necessarily valuable, involves casting aside or denigrating what is contingently valuable, and this in turn includes many elements of a well-lived life, like one's personal projects.

There is thus an idea shared between the psychological realist critique of modern morality and my concern with the existential guilt of depressed persons: an ethical life should also be a well-lived life and should not demand otherwise. I would instantiate this

156

general notion as a single claim: it should be possible and permissible for an agent to act simply because they want to, not because they must. Williams writes that "it is a mistake of morality to try to make everything into obligations" (180). This is a mistake of the depressed, too.

Yet the depressed person, rather than being alienated from morality, seems to compulsively flee *into* morality. If what I have said thus far about depression and motivation is correct, then rather than the problem being that morality is alienating to depressed persons, the issue is instead that depressed persons engage with moral obligations as already alienated persons unable to pursue their desires when entrapped by existential guilt. Certainly, there are cases where some kinds of ethical conduct can be more difficult for depressed persons, like keeping promises or having self-respect. But understanding and sympathy for depressed persons, while far from universal, is increasingly common. As long as depressed persons are seen as unable, and not unwilling, to act, even the strictest ethical theory would not claim them to be responsible for said inaction. The case of depressed persons might thus seem to intersect only superficially with the arguments of psychological realism, especially since depressive alienation likely has causes entirely unrelated to the structure of modern morality, and especially given that depressive guilt distorts one's conception of what morality demands. But depressed persons diverge from the psychological realist's vision of moral agents in an interesting way. The depressed person appears to be alienated not so much from their obligations as from their own desires, struggling less to care for others than to care for themselves. It seems that for a depressed person, it is prudence, not morality, that is too demanding.

157

There are, I think, two ways to understand this inversion of the psychological realist critique of excessive demandingness. One option is that a depressed person distinguishes moral obligations from prudential concerns and compulsively devalues the latter. But I think a phenomenologically more accurate account is that, for the deliberating depressed person, the *prudential becomes moralized*: all prudential concerns show up within a broader web of moral concerns such that the former inevitably appear unworthy in light of the latter. Every possible activity is viewed within a field of guilt and self-blame; all action seems to require the possibility of self-forgiveness for failing to act better. This amounts to a short-circuiting of ethical deliberation that inhibits the depressive person in reflectively endorsing a personal hierarchy of values or goals. Moreover, the specific way in which depressed persons become alienated from their prudential concerns parallels a psychological realist critique of utilitarianism: the demand to evaluate all action in moral terms will elevate moral acts over prudential self-interest to an extent that denies and destroys one's personal identity and distinctness as an individual moral agent. Seen thusly, depressed persons enact compulsively what at least some strict theories, like Peter Singer's version of utilitarianism, require rationally. This comparison should not be taken too far. Depression hardly facilitates action, so this compulsive inability to maintain a moral-prudential distinction would not make a depressed person a "good" utilitarian, at least in practice. But if it is true that depressed persons cannot consistently maintain a moral-prudential distinction in deliberation, any strict theory that necessarily involves such a distinction would be ill-suited to depressed persons even though they are otherwise rational agents. The case of depressed persons thus generates a new argument in the vein of psychological realism: if a strict theory

158

requires a depressed person to, say, recognize that moral obligations have a special status distinct from prudential concerns, that theory is asking too much.

I present a new psychological realist argument in favor of a moderate or "reliable" ethics and against a strict ethics: depressed persons compulsively efface distinctions between the moral and the prudential in deliberation, presenting a challenge to any theory of moral obligation that assigns it a special status distinct from the prudential. Any general "overridingness" of moral obligations over prudential concerns will, if it is to be put into practice, likely depend on being able to distinguish obligations with this special moral status from other obligations or interests. But while the depressed person is demoralized with regard to motivation, they are "hypermoralized" with regard to deliberation: every possible act seems to have an ethical status, every act not taken seems to deepen their perceived ethical unworthiness, and every act taken seems to fall short of a vague but absolute ethical ideal.

I further argue that what specifically is weakened in hypermoralized deliberation, and what leads in practice to the effacing of a moral-prudential distinction, is a "sense of agency" in the experience of forming intentions. The experience of sensing one's own agency at work in intention-formation plays an epistemic role in ethical deliberation that has been insufficiently examined. This experience is suppressed by the sense of having already and inescapably "fallen" into ethical unworthiness that I have associated with existential guilt, but not completely. If exercising the capacity of ethical deliberation is good, then cultivating this sense of agency is also good insofar as it contributes to this capacity. And if a moderate ethical theory aligned with some form of psychological

159

realism emphasizes the contingently and personally valuable, it may in practice cultivate the sense of oneself as an agent, a distinct person with one's own value set.

In my original discussion of Cholbi's account of depressive listlessness in Chapter 1, I modify his use of the terms "moral" and "prudential" into other-concern and self-concern, due to the possible complication of self-respect being a moral obligation rather than a prudential interest and Cholbi's focus being altruism. This modification is also useful in a decision procedure for an ethics for the depressed. That is, what a depressed person best attends to in ethical deliberation is not a distinction between moral obligations and prudential interests but between "other-regard" and "self-disregard." These are intuitions that may show up in a compulsive tangle in the context of existential guilt: for example, "I must show more regard for others because I must redeem my unworthiness"; and "I must redeem my unworthiness because I show insufficient regard for others." For a depressed person to avoid having deliberation spiral into rumination or hypermoralized deliberation, they benefit from cultivating habits or dispositions to 1) attend to the difference between other-regard and self-disregard within this tangle of intuitions and 2) trust the other-regarding elements of intuitions and distrust the self-disregarding elements of intuitions. To take existential guilt as the starting point for a reliable ethics is, in a sense, to take *half* of existential guilt as a starting point. But existential guilt does not manifest conveniently divided into halves. Intuitive or compulsive other-regard and self-disregard become more easily distinguishable for the depressed through treatment and the cultivation of mental habits or techniques.

To identify hypermoralization with the loss of a sense of agency in intention formulation is not just to say that depression undermines planning as well as action. It is

160

also to emphasize the point of psychological realists like Williams: the experienced expansion of the domain of what is moral is also the experienced retreat of the domain of what is mine. Shaun Gallagher distinguishes between a sense of agency and a sense of ownership with regard to physical movement. But perhaps, as Gallagher's research with Dylan Trigg suggests, in intention formation and the phenomenon of "mineness" in deliberation, this distinction between agency and ownership is less phenomenologically clear. A loss of felt or intuited ownership may seem also like a loss of felt or intuited agency over one's own thinking.

One objection to my line of thinking, commonly raised in defense of utilitarianism, is that a moral theory is not necessarily a decision procedure: even if a depressed person is supposed to value the moral above the prudential in some sense, this does not mean that they must think in those terms. This resonates with the major "therapy, not theory" objection to my project that I discussed in Chapter 1: if a depressed person's demoralization is not going to be responsive to theoretical arguments, it seems both mistaken to find fault in the structure of ethical theories and mistaken to respond by modifying said theories. Here, because the issue lies in a depressed person attempting to put an ethical theory into practice in deliberation and being unable to maintain a moral-prudential distinction arguably necessary for some strict theories, it is important to distinguish between the theory itself and how to put the theory into practice. It might be that an ethical theory is best explained or justified in terms of a moral-prudential distinction but that one might affirm and act in accord with that theory without, when practically deliberating, thinking in the terms by which the theory is explained or justified. If this is the case, then a formalist or procedural ethics might accommodate a

depressed agent by developing a decision procedure that does not straightforwardly rely on knowingly assigning some obligations a special status even if, functionally, the agent does so.

My response to this objection is tentative: I think a depressed person will also struggle to maintain in practice the distinction between moral theories and decision procedures, that is, be inclined to turn away from the procedure toward the theory that justifies the procedure. This person seeks justificatory reasons out of a desperate desire for "more motivation," that is, the capacity to act. The philosophical nuts-and-bolts will matter to them not just theoretically but practically, because the ruminating depressed person will be inclined to reenact the justification of the decision procedure when struggling to act. This is another way in which seeking a mode of ethical thinking tailored to depressed persons may matter: they tend to read the theory. I may be overgeneralizing. But if there is anything to this idea, and if there is a form of ethical thinking that might be liberating to depressed persons despite the argument-resistant quality of their compulsions, then it may also be worth formulating this way of thinking as a theory.

To make the case for depressive psychological realism regarding the moral-prudential distinction in deliberation, I will begin by reviewing the literature on psychological realism more broadly, focusing on Williams's initial arguments, Owen Flanagan's critique of Williams, and David Wong's overview and assessment. I will then discuss Shaun Gallagher's research into the sense of agency in intention-formation and how it applies in the context of ethical deliberation. Finally, I will consider alternatives to a strict moral-prudential distinction as discussed by Julia Annas and Samuel Scheffler and, returning to Wong's discussion of Scheffler, argue that a moderate ethical theory can

162

foster a depressed person's sense of agency by not relying on the moral-prudential distinction operative in assigning moral obligations a special status. Such a moderate theory thus has an advantage over strict theories that undervalue the contingent in attempting to secure a necessary value for the moral. However, a depressed person's compulsive strictness toward themselves might lead them to balk initially at a moderate theory and its implicit acceptance of some level of moral luck. I will thus proceed in the next chapter to develop a normative ethics for the depressed that I hope will be acceptable both to depressed persons and to ethical theorists: a value ethics according to which values are felt intuitions and norms are fundamentally predictions about what will realize these values broadly and in the long term.

1. Williams and Flanagan on Psychological Realism

The phrase "psychological realism" implies that morality decoupled from it is unrealistic, and for Williams, though there are many symptoms of this decoupling, one unrealistic ambition is its central cause: the attempt to eliminate the threat posed to moral authority by moral luck. Williams's critiques of what he calls the morality system range widely but focus on the theme of privileging necessity over contingency. The morality system, Williams claims, places too much emphasis on obligation as the primary or sole relevant moral consideration such that other considerations that seem intuitively to be closely tied to moral decision making, like regret, are declared irrelevant. Practical necessity, the link between rational deliberation and moral action, becomes understood exclusively, and thus falsely, in terms of obligation and morality. The notion of a purely voluntary act becomes essential as what is involuntary becomes morally irrelevant and

163

the assignment of blame for what is voluntarily chosen becomes absolute in the pursuit of a pure ideal of justice. But moral luck, broadly meaning the notion that there is some form of chance or randomness to how much one is to be credited or blamed for one's actions or character, implies that there is a limit to how fair morality itself can be in determining what assignations of credit and blame are justified. To say that moral luck is real is to say that it is impossible to design a coherent moral theory that only credits or blames an individual for what they are really responsible and never credits or blames them for, say, circumstances beyond their control. Yet if a moral theory cannot provide the necessary criteria for so assigning credit and blame, then its authority is no longer the authority of reason itself: any moral agent could reasonably reject this theory's dictates or evaluations of moral status without, say, necessarily self-contradicting. Williams's cluster of critiques of the morality system are best understood as diagnosing the symptoms of the rejection of moral luck and prescribing some acceptance or acknowledgment of moral luck's reality as the cure, or at least the start of one.

Although Williams explicitly rejects the use of philosophical argument as a response to despair, as I noted in Chapter 1, my reading of his work focuses on the parallel between his concerns with the morality system and what I take to be a depressed person's: if morality imputes to us a kind or degree of agency that we lack and then blames us for not having it, it is unfair and unreasonable. Moreover, if the goal of developing a strict moral theory is to eliminate the unfairness of moral luck, then to reproduce or deepen that meta-level unfairness by falsely framing contingent assignations of credit or blame as absolutely authoritative is to fail on its own terms. A moderate theory, one that is hesitant or hedging in how it assigns credit or blame, might be

criticized for tolerating hypocrisy or "moral holidays" from moral agents in a way that is also unfair to those who are properly moral, thereby failing to consistently assign credit and blame in accord with the values the theory ostensibly promotes. But its defense would be that this is the most fairness one can expect from a moral theory, that is, conceding that the theory is not absolutely authoritative. Framing Williams's ethics thusly resonates with the challenges of depression more than a reading that focuses on his integrity critique of utilitarianism.[34]

I pause to emphasize a limit of this parallel with depression: I assume that depression and its attendant losses of motivation are broadly understood as not being a depressed person's fault, and I further assume that most or all strict moral theories are compatible with this understanding. If a depressed person has diminished agency, a strict moral theory that aims to apportion credit and blame with absolute correctness and consistency could acknowledge that diminished agency as relevant to said apportioning, thus blaming depressed persons less or not at all for, say, lacking certain virtuous traits like punctuality or proper self-respect. The challenge posed by depression to the morality system as Williams describes it is subtler and deeper than a simple rejection of misplaced blame. First, the demoralization of the contingent, the stripping of moral value from what

---

[34] That said, Williams's integrity critique, that being that utilitarianism ignores and denies the separateness of moral agents by obviating the value of their personal commitments and projects, draws on his discussion of personal identity, particularly the problem of self-persistence and the possibility of justifying present action for the sake of one's future self. These challenges resonate with Ratcliffe's account of how a depressed person's experience of time is compressed and distorted. One might frame the depressed person's struggle to deliberate as an inability to intuit with sufficient motivational force that their future self is their present self persisting, then frame this deliberation process as an attempt to justify claims to oneself that depend on identifying with one's future self. This would draw on a broader swath of Williams's work, but risks conflating problems of justification with losses of motivation: justifying claims "to oneself," in the case of a depressed person ruminating, likely has little in common with justifying claims to another rational agent and much more in common with repeatedly running face-first into a brick wall. I thus focus primarily on the themes of moral luck and less so on integrity in this chapter.

is not necessarily indicative of how credit and blame ought to be properly apportioned –
like, say, regrets that may or may not indicate an actual moral error – is itself a depressive
move. The search for an ultimate grounding or justification for moral authority
unsettlingly resembles the desperate depressive search for a motivation that depression
cannot defeat. Second, as a depressed person's motivations become demoralized, so do a
depressed person's deliberations become "hypermoralized": existential guilt takes every
perceived moral obligation or opportunity and renders them a reflection of the agent's
seeming moral inadequacy, effacing any distinction between genuine moral concerns and
merely prudential interests. I think the relevance of these points depends on the rightness
of my earlier suggestion that a depressed person tends compulsively to blur the
distinction between a theory and a decision procedure, preventing a strict moral theory
from developing some gentler means of putting itself into practice and thereby shielding
a depressed person from its own strictness. But if this is true, then for a depressed person,
an ethical theory is likely to slip into becoming a way of ethical life. And as a way of
ethical life, a strict theory both fosters a depressed person's most self-destructive
compulsions and, in devaluing the prudential through a distinction that a depressed
person struggles in practice to maintain, fails to affirm their ethical worth. If Williams is
right that an aversion to moral luck is behind the failings of strict moral theories, and thus
that a moderate approach to ethics is the solution, the question becomes what a moderate
theory entails and why.

By beginning with Williams's initial critique of the morality system and then
turning to Owen Flanagan's modified Minimal Psychological Realism and David Wong's
normative incorporation of the position in the form of his pluralistic relativism, I aim to

166

build toward the proposition that a moderate ethical theory may be compatible with both (1) a broad sense of what is possible for and (2) broad restrictions on what constitutes an "adequate" ethical theory, but *not* the claim that (3) individual flourishing requires living an ethical life. There would be no ultimate justification at the individual level for living an ethical life, say, one that renders it the necessary consequence of self-consistent rationality, but there would be space to take seriously – that is, to take as central and proper subjects of ethical theory and matters of ethical life, and not as the products of some sort of failing – substantive moral disagreement, as is Wong's primary concern, and the practical inevitability of internal value conflict, as will be mine.

In the first chapter of *Moral Luck* (1981), Williams extends his integrity critique of utilitarianism to Kantianism and what he calls "impartial morality" generally. Though utilitarianism and Kantianism evaluate moral acts based on a different locus, consequences and intentions respectively, and so may generate different moral conclusions about acts like lying, both are concerned with producing "the impartial good ordering of the world of moral agents" (14). The procedures with which they are associated, involving the application of the Greatest Happiness Principle or the varied formulations of the categorical imperative, are designed to be indifferent in their assignment of obligations or blame to contingent factors like an agent's temperament or desires. Williams argues that this impartial approach to morality ignores a category of commitment he calls a "ground project": there are pursuits so central to a person's existence that their personhood depends on, at the very least, taking the pursuit as relevant to moral decision-making, that is, as potentially a reason to choose against what the procedure dictates is the moral choice. He writes:

somewhere … one reaches the necessity that such things as deep attachments to other persons will express themselves in the world in ways which cannot at the same time embody the impartial view, and that they also run the risk of offending against it. They run that risk if they exist at all; yet unless such things exist, there will not be enough substance or conviction in a man's life to compel his allegiance to life itself. Life has to have substance if anything is to have sense, including adherence to the impartial system; but if it has substance, then it cannot grant supreme importance to the impartial system, and that system's hold on it will be, at the limit, insecure (18).

The insecurity at the limit of morality is that for the notion of a singular moral agent to be coherent, this agent requires a singular identity. But if that identity depends on a ground project that gives life "substance" but may or may not be a moral good consistent with the impartial system, that system 1) may produce tensions between what is good insofar as it is a condition of agency, and thus any good moral decision whatsoever, and what is supposed to be good according to the system's principles and 2) thereby introduces partiality to what is good for different agents with different ground projects, ceasing to be entirely impartial. Even if, for Williams, Kantianism does not deny the separateness of agents as straightforwardly as utilitarianism, insofar as it does not take ground projects seriously as, e.g., even potentially important enough to override the requirements of the categorical imperative for a particular agent, the same insecurity arises. The problem seems to the ambition to impartiality itself.

This problem becomes acute in his next chapter and the context of moral deliberation: only if one identifies with one's future self (or selves, perhaps) is it possible to get a sense of what might amount to a justified or even rational decision. While moral luck can refer to random chance interfering with a specific act, its more troubling version is how we are constrained by circumstances from the moment we are born, to the point that identifying in an agent the singular sovereign will making moral decisions can seem

168

implausible. "One's history as an agent," Williams writes, "is a web in which anything that is the product of the will is surrounded and held up and partly formed by things that are not" (29). By linking personal identity with a ground project, Williams wears away at the foundations of impartial morality in the conception of the moral agent as defined by purely voluntary acts.

One wedge issue and recurring example for him is regret. Whether one is responsible for the consequences of or intentions behind one's actions, an impartial morality attains impartiality by restricting responsibility to what (respectively) follows from or motivates one's choices. It seems to follow that while remorse for one's immoral voluntary acts is both rational and a properly moral matter, regret over, say, unforeseeable events or simply what might have been is, if not irrational, at least not a "moral" sentiment in the same respect insofar as it does not pertain directly to voluntary acts. Williams concedes that that someone can rationally conclude both that they regret something and still would not act otherwise because "he stands by the processes of rational deliberation which led to what he did" (32). But those processes of rational deliberation are dependent on a standpoint itself grounded, Williams thinks, in a ground project. There are projects that one might regret or not and then there are projects – ground projects – that at least in part define what sort of projects may count as regrettable.

To relegate both sorts of projects to "nonmoral" status insofar as both are contingent, arising from whatever deep attachments an agent happened to acquire, thus undermines the possibility of assigning "sentiments or attachments," intuitively serious

matters of ethical life, any moral weight whatsoever. Williams concludes (I quote at length):

> The ultimate justice which the Kantian outlook so compellingly demands requires morality, as immune to luck, to be supreme, and while that does not formally require that there be no other sentiments or attachments, in fact it can, like the Robespierrean government to which Heine compared the Kantian system in general, steadily grow to require a wider conformity of the sentiments. Justice requires not merely that *something I am* should be beyond luck, but that *what I most fundamentally am* should be so, and, in the light of that, admiration or liking or even enjoyment of the happy manifestations of luck can seem to be treachery to moral worth. That guilty levelling of the sentiments can occur even if one recognizes, as Kant recognized, that there are some things that one is responsible for, and others for which one is not. The final destruction occurs when the Kantian sense of justice is joined to a Utilitarian conception of negative responsibility, and one is left, at any level of importance, only with purely moral motivations and no limit to their application (38).

His incidental invocation of "guilt" in "guilty levelling of the sentiments" suggests another minor parallel between Williams's ethics and an ethics for the depressed. The role that regret plays for Williams seems similar to the role of existential guilt in this project: it is a kind of lingering remainder for a strict ethics, central to the lived moral experience of an agent yet relegated by the theory to nonmoral status or irrational detritus, even when this restrictive conception of agency may risk making a well-lived ethical life unimaginable or impossible.

In the final chapter of *Ethics and the Limits of Philosophy* (2006), Williams presents a further consequence of the strict understanding of deliberation: the focus on obligation within the impartial "morality system" as the primary sort of moral consideration. Moral obligation is for the morality system a special category of consideration distinct from obligations generally. Williams identifies four aspects of moral obligation: particular obligations are supposed to follow from more general obligations; moral obligations are practical conclusions and so "ought implies can" in the

sense that, if the proposed obligation resulting from moral deliberation is impossible for the agent, the agent must deliberate again; moral obligations must ultimately cohere with each other; and moral obligation is "inescapable" or categorical (177). On the last point, moral obligation is inescapable both in the first-personal sense that it "applies to people even if they do not want it to" and in the third-personal sense that it "can apply to people even if, at the limit, they want to live outside that system altogether," or in other words, insofar as one is a "responsible agent," there is "nowhere outside the system" (178). The morality system must bind all, in the sense that no one can simply opt out of being rightly morally judged: even if someone utterly rejects morality itself, it must remain true to say that they *should* behave according to it, and their rejection must be irrelevant to that truth.

But Williams argues that taking this special sense of obligation as the primary kind of moral consideration slips into taking it as effectively the only kind of moral consideration, because no other sort of consideration, like sentiments or attachments, can supersede it. The first aspect of moral obligation, that particular obligations follow from more general ones, emerges because only obligations can overcome other obligations:

> It is a mistake of morality to try to make everything into obligations. But the reasons for the mistake go deep. … [Moral] obligations have a moral stringency, which means that breaking them attracts blame. The only thing that can be counted on to cancel this, within the economy of morality, is that the rival action should represent another and more stringent obligation. Morality encourages the idea, *only an obligation can beat an obligation*. (180)

Williams calls this the "obligation-out, obligation-in" principle and suggests that two main troubles follow from it. First, it seems to follow that moral obligations culminate in a single general and ultimate moral obligation, but it is difficult to express very general obligations *as* obligations, suggesting that "obligation" may not be the best conceptual framework for understanding an agent's most fundamental relation to ethics or reason for

171

being ethical. Second, what Williams previously referred to as "negative responsibility"

emerges: "if we have accepted general and indeterminate obligations to further various

moral objectives, they will be waiting to provide work for idle hands," he observes, and

"the thought can gain a footing (I am not saying that it has to)" that "I am under an

obligation not to waste time doing things I am under no obligation to do … If obligation

is allowed to structure ethical thought, there are several natural ways in which it can

come to dominate life altogether" (181-182). The notion of "a duty to myself," which

Williams takes to be "fraudulent," arises precisely because in order for one to act in one's

own interests when bound by moral obligation, one must have a moral obligation to do

so. Despite such polemical language, Williams seems hesitant to say that moral

obligation is necessarily problematic or dominates a moral agent, writing in terms of

predictions of how it might shape one's ethical life (and not saying that it has to). Even

so, he takes obligation to be the morality system's central cog, turning to the question of

how to better incorporate some modified notion of obligation as one significant ethical

consideration among many.

Williams moderates the status of obligation as a moral consideration by giving it a

mediating role between the broad notion of "importance" and the narrower function of

deliberative priority. The "notion of *importance*," he says, is what will "help to lead us

away from morality's special notion of moral obligation, and eventually out of the

morality system altogether" (182). Rather than define "importance" outright, Williams

instead takes pains to avoid a reductive approach: when some ends, like "not being killed

or used as a resource," are important, "*one* way in which ethical life serves [those ends] is

by encouraging certain motivations, and *one* form of this is to instill a disposition to give

the relevant considerations a high deliberative priority." It does not seem important, as it were, to understand every dimension of what "importance" might mean in order to identify the forms of importance, like goals and purposes, that become relevant for moral deliberation downstream. In this context, obligations amount to "considerations that are given deliberative priority in order to secure reliability" (185). He summarizes his version of moral obligation pragmatically, as follows:

> Obligation works to secure reliability, a state of affairs in which people can reasonably expect others to behave in some ways and not in others. It is only one among other ethical ways of doing this. It is one that tries to produce an expectation *that* through an expectation *of*. These kinds of obligation very often command the highest deliberative priority and also present themselves as important—in the case of promises, because they are promises and not simply because of their content. However, we can also see how they need not always command the highest priority, even in ethically well-disposed agents. … We should reject morality's other maxim, that only an obligation can beat an obligation. (187)

More subtly, he works via this redefinition to detach practical necessity from obligation: "the course of action the agent 'must' take may not be associated with others' expectations, or with blame for failure" (188). Here the idea of integrity or a ground project returns. While the "necessity" of practical necessity implies that it cannot follow from a contingent factor, like desires generally, "if the desire were not one that the agent merely happened to have, but was essential to the agent and had to be satisfied," then a course of action that "must" follow could follow from said desire (189). The weight of binding morality to rationality itself, as a means of deriving absolute authority for morality through non-self-contradiction, is lifted from the notion of obligation so long as a ground project may operate as a contingency that acquires the force of a necessity. It is one's contingent ground, not one's rational nature, that justifies the basic idea of practical necessity as a bridge between thought and action: there are matters where if one truly

173

believes X, one *must* act accordingly. These matters are likely to change along with this justification. To betray one's highest values or purposes, whatever those happen to be and even if they are not generalizable (let alone universalizable) goods, becomes by this account a deeper kind of self-repudiation than breaking a promise or telling a lie. This picture of ethical life is, I think, Williams at his most intuitively correct.

Finally, with the shift from necessity to consistent contingency comes a shift from moral agency as an individualized to a more thoroughgoingly social phenomenon. Williams thinks that the morality system obscures its own sociality:

> Since ethical considerations are in question, the agent's conclusions will not usually be solitary or unsupported, because they are part of an ethical life that is to an important degree shared with others. In this respect, the morality system itself, with its emphasis on the "purely moral" and personal sentiments of guilt and self-reproach, actually conceals the dimension in which ethical life lies outside the individual. When we know what the recognition of obligation is, if we still make it the special center of ethical experience, we are building ethical life around an illusion. (191)

The illusion that practical necessity can only follow from the special moral mode of obligation fosters a kind of moral myth of voluntarism, that it is both possible and desirable to ascertain with absolute certainty what individual choices are "real" choices issuing from one's pure and sovereign will, guided by rationality and unimpeded by sentiment or attachment. This absolute certainty would be required to assign responsibility, credit, and blame with absolute impartiality. The impossibility of acquiring this certainty coupled with the essential need for it transforms the morality system into a "blame system" that spurns indirect forms of moral guidance:

> There is a pressure within [morality] to require a voluntariness that will be total and will cut through character and psychological or social determination, and allocate blame and responsibility on the ultimately fair basis of the agent's own contribution, no more and no less. It is an illusion to suppose that this demand can be met … This fact is known to almost everyone, and it is hard to see a long

174

future for a system committed to denying it. But so long as morality itself remains, there is danger in admitting the fact, since the system itself leaves us, as the only contrast to rational blame, forms of persuasion it refuses to distinguish in spirit from force and constraint. (194)

For a morality system that derives its impartiality from its understanding of moral agency as pure and absolute, "forms of persuasion" like nudges, presenting some options over others without appealing to the agent's rationality directly, are at best nonmoral or morally neutral or at worst insidious deprivations of opportunities to properly choose. This amounts to a vision of morality in which the social situatedness of the moral agent is an incidental fact, not one essential to the possibility of deliberating coherently in the first place. Of course, one can hold (as Williams does in his discussion of confidence) that appealing to an agent's rationality first and foremost is generally for the best and that a society instead dedicating itself to subconsciously manipulating its members into behaving "more ethically" would at best set dangerous precedents. But, as he states, although the latter is extreme, it is an extreme on a spectrum of what ethical life always and already involves, that is, environmental cues. To exclude such indirect methods from the proper governance of ethical life is not to protect agency but to impoverish it.

To accept the reality of moral luck, he concludes in his final paragraph, is thus to revitalize ethical life:

> Many philosophical mistakes are woven into morality. It misunderstands obligations, not seeing how they form just one type of ethical consideration. It misunderstands practical necessity, thinking it peculiar to the ethical. It misunderstands ethical practical necessity, thinking it peculiar to obligations. Beyond all this, morality makes people think that, without its very special obligation, there is only inclination; without its utter voluntariness, there is only force; without its ultimately pure justice, there is no justice. Its philosophical errors are only the most abstract expressions of a deeply rooted and still powerful misconception of life. (196)

Psychological realism rejects impartial morality not because it is overly demanding in the content of its moral requirements, then, but because its pursuit of an impossible confidence in one's moral conclusions is destructive to the profuse variety of ethical life, rich with considerations that are all worthy of theoretical legitimacy as "moral." Introducing partiality to the heart of ethics in the form of an individual's ground project does mean that, even at their most coherent, reasonable moral conclusions will contradict one another when individuals with different ground projects come into conflict. This way of ethical thinking may even provide, in practice, cover for bad actors appealing to their own integrity to justify their excesses or to hypocrites who declare themselves exempt from inconvenient obligations on the basis of the ground project they claim to have. But it is a way of ethical thinking that can go on in the face of moral luck and need not live in denial, one that is open to the diversity of ways in which persons really do become ethical. It is a way of ethical thinking that is, in these senses, realistic.

That said, the major vulnerability of Williams's conception of ethics is the centrality of the "ground project" to both the form and content of ethics, in the sense that practical necessity, the deliberative priority of ethical considerations, and the particulars of one's integrity alike follow from it. With so many key features of ethical life now seeming to depend on the presence of a ground project, although the contents of that project might be contingent, an individual having such a project (or projects, or deep attachments, or whatever stands in) no longer seems optional if we are to think of them as having a recognizable ethical life. This may be too narrow an understanding of our psychology: it is not only conceivably possible that a moral agent could lack a ground

176

project or deep attachments while remaining, at least in some minimal or underdeveloped but still recognizable sense, an agent, but it is plausible that many actually do.

Owen Flanagan's minimal psychological realism adapts and clarifies Williams's theory while presenting two main challenges to the concept and role of the ground project. First, Flanagan sees a philosophical provincialism in Williams's radical rejection of impartiality: instead of assuming that Kantianism and utilitarianism are the dominant ethical alternatives, one ought to incorporate in psychological realism the possibility of personal commitments to impersonal philosophical positions such as a Buddhist relinquishment of desire. A more worldly or comparative view of ethics might not only reveal a variety of ways of valuing impartiality compatible with the acceptance of partiality at the heart of ethics, but also loosen the notion of a "ground project" to better include community-oriented commitments. Second, and I think more fundamentally challenging, Flanagan emphasizes that a truly social understanding of ethics (for which Williams appears to aim) accepts, basically, that integrity can be bad and alienation can be good depending on the content of one's commitments, and that endorsing this view is essential to any vision of social progress. Williams has a pluralistic approach to what counts as a legitimate ethical consideration, and he is more concerned with what conclusions count as justified than what is "really" good. But it may still be going too far, or romanticizing the individual's inner struggle or existential quest too much, to say that a project or attachment may have an ultimate deliberative priority or the force of practical necessity. Flanagan's concern for a social ethics motivates a critique of Williams's fixation on the so-called morality system, one that parallels the "therapy, not theory" objection to this work: Williams weaves between discussions of ethical theory and ethical

177

life because, in the end, what he wants to criticize is a particular society, and to criticize a

society by way of criticizing a theory is to aim shy of one's mark and, in so doing,

demand too much of theories. Arguably, even as Williams is vigilant not to presume that

to change theory will be to change practices, his construction of "the morality system" as

his nemesis distracts him from offering a more concrete social critique of the self-

undermining ethical life that "the morality system" represents in the abstract.

In *Varieties of Moral Personality* (1991), Flanagan argues for the relevance of

moral psychology to moral theory through his Principle of Minimal Psychological

Realism (PMPR): "Make sure when constructing a moral theory or projecting a moral

ideal that the character, decision processing, and behavior prescribed are possible, or are

perceived to be possible, for creatures like us" (32). Flanagan's work marks a significant

effort to expand moral psychology beyond questions of behavior or motivation to

personality formation or varying conceptions of virtues. He grants that "knowledge of

local personality organization, of what is considered natural, expectable, and mature in a

certain vicinity, can never settle by itself questions about what is good" (16). But:

> even a regulative morality will draw on an image of ourselves that we are capable
> of admiring and to which we can in some sense imagine conforming. A normative
> conception which fails to meet certain standards of psychological realizability will
> fail to grip us, and in failing to grip us will fail to gain our attention, respect, and
> effort. (26)

By integrating empirical and philosophical psychology and taking "ideals of human

personality which incorporate more than motivationally necessary conditions for social

harmony" (19) to be ethically relevant, Flanagan aims to determine what these standards

of psychological realizability might be, consonant with a modern understanding that

"deep and universal facts are relatively few and far between, and … socially embedded

generalizations are all subject to displacement with changes in the social structures that make them true" (334).

He contrasts his minimal standards with what he calls "strong realism," which asserts that "modern moral theory is too demanding" (46) and of which he takes Williams to be a proponent. Flanagan, citing Williams and Susan Wolf, takes the general strategy of strong realism to be to claim that the separateness and distinctiveness of individuals justifies the "existential significance of unique life plans and ground projects" (59), limiting to what extent an ethical theory may require impartiality, alienation, or violations of integrity. But, drawing on Samuel Scheffler (1982), Flanagan emphasizes both the capacity of individuals to "override their natural partiality" (71) and to commit themselves personally to "an extremely impersonal moral perspective" (70) like a Buddhist ideal of relinquishment. Even if ground projects are a general feature of human psychology, neither their generality nor their developmental significance necessarily provide a basis to object to the content of specific ethical theories. What Flanagan objects to in strong realism is the "underlying presumption" that "the identity conditions of persons—the personality, projects, commitments, and so on which make them the persons they are—set some sort of deep and inviolable constraint on what an ethical conception can demand." He continues:

> The only way of gaining this premise which I have seen attempted in the literature—beyond some intuition pumping by Wolf about certain widespread contemporary attitudes about privacy and fun—is the route via Williams' argument … to the effect that such identity conditions give each life whatever meaning it has. The trouble is that the argument which yields this conclusion (and possibly also the conclusion that every person has "internal reason" to want to achieve *whatever* will make his life meaningful) does not also and at the same time yield the conclusion that a life rendered meaningful by its identity conditions is a life worth living or, what is different still, a life whose shape and substance ought to be respected by any credible ethical theory. (82)

Flanagan seems to attribute Williams's attempt to delimit the scope of credible ethical theories to a misdiagnosis of a communal problem. "It falls on a failure of the moral community," he argues, and "not on the degree of abstraction and impartiality" of a theory or conception, "to equip persons with the self-management skills to avoid developing deeply incompatible commitments, and with the ability to override such incompatibilities in the desired direction if they do occur" (91). Better to cultivate practical wisdom, then, than to reject impersonal ethics.

I think that while Flanagan is right to be dubious of how much ethical content, like a normative commitment to integrity and against alienation, can be justifiably derived from the theory of persons undergirding psychological realism, he is to an extent talking past Williams in his criticisms precisely because he takes Williams's pluralism for granted. Flanagan interprets moral theories through a pluralistic lens, presenting a range of contradictory moral theories as all credible and legitimate, even if those theories are stringent and would themselves reject his pluralism. Williams, however, takes that pluralism to be a theory in need of defense against moral theories that would reject it. It might be true that an individual could employ the categorical imperative or the Greatest Happiness Principle, even commit themselves to doing so consistently, and rely on some capacity of practical wisdom to avoid compromising other commitments that must never overrule said procedures, like never forming an attachment deep enough to threaten the theory. If such an individual deserves to be called a fully rational person, then strong realism has indeed been refuted, and insofar as that is Williams's stated position, so has he. But it is not clear to me that the imperative to engage in such contortions, by developing one's personality in such a way that one may act self-consistently with an

180

impersonal morality, is itself consistent with the impartial morality system that Williams criticizes, for which personality formation is a secondary moral concern insofar as it contingently facilitates necessary acts.

Moreover, if it is true that impartial morality aims to enable certainty and eliminate luck in moral conclusions, then even if some capacity to perceive when a moral situation has arisen is always required in the application of procedural morality, part of the point of a procedure is to reduce the need in moral deliberation for something so nebulous as practical wisdom. Flanagan's pluralistic vision of potential moralities justifies both his generosity toward impersonal ethics and his confidence in practical wisdom, but his version of impersonal ethics is already personalized, presented as the product of a personal commitment. If Flanagan is right, then impersonal ethical theories work because they are wrong about themselves as necessary, certain, and overriding, relying on extensive personality formation which obviates the need of the procedure to override serious sentiments or attachments just to appear to succeed. Paradoxically, Flanagan can claim that the pursuit of eliminating moral luck is rational as, say, a regulative ideal, but the pluralistic grounds on which he would justify this – especially his core emphasis on personality formation as a primary feature of ethics – mean he cannot self-consistently claim that eliminating moral luck is possible. To my mind, he has already conceded everything that Williams would want from psychological realism.

Even if I am right, Flanagan makes a compelling case for not universalizing a ground project as a central feature of moral psychology as Williams does, and the question becomes what a refined normative project incorporating psychological realism looks like. Flanagan himself does not have a normative ethical project, strictly speaking,

181

so much as a descriptive project with some consistent guiding normative goals.[35] The

best example of such a project is David Wong's *Natural Moralities: A Defense of*

Pluralistic Relativism (2006), a nuanced account that proposes universal constraints on

moralities based on necessary conditions for flourishing while including a wide range of

communities as having adequate moralities and thus affirming the possibility of

substantive and serious moral disagreement. Given his primary concern with

intracommunal and intercommunal disagreement, Wong offers only a limited discussion

of moral considerations or deliberation at the level of the individual. Even so, it is during

this discussion that he makes a striking claim from the perspective of an ethics for the

depressed: there is no ultimate justification for living morally. Wong rejects the claim

that individual flourishing might require a moral life, arguing instead that all a moral

philosophy can justify is the moderate claim that individual flourishing and a moral life

are potentially compatible. No morality can, then, justifiably settle in advance a choice

---

[35] In *The Geography of Morals* (2016), Flanagan summarizes the effect of descriptive ethics on normative ethics as indirect but significant and castigates the arrogance and narrowness of moral philosophers who default to centering their own presumptively normative language and intuitions, especially given that empirical moral psychology remains a young science. "If normative ethics is to be helpful in the project of living well, of flourishing, of finding meaning and purpose, of leaving the world a better place," he concludes, "it ought to help us to be attentive, sensitive, and open to value, not cocky, overconfident and closed to other ways of thinking and being. … Knowing where to go from here, how to go on, what to do next requires knowing what the possibilities are" (15). The constraints that Flanagan sets on normative ethics are mild, recommending a focus on persons within communities rather than abstracted from them, a nonreductive view of human nature conscious of the variety of moral personality, and cross-traditional inquiry (21). But though he expresses great annoyance with moral philosophers who dismiss descriptive ethics as beneath the concerns of normative ethicists just because "ought" does not follow from "is," he is, interestingly, also suspicious of provincialism from those who claim to naturalize ethics, like some evolutionary psychologists, but blur the distinction between the descriptive and the normative such that they naturalize (and thereby normatively entrench) contingent moral factors. Flanagan rejects the view that "substantive moral criticism and positive individual change are weak forces, and relatedly … the view that the distinction between the descriptive and the normative is a thin and uninteresting one, that people pretty much are as they ought to be, given that the sum of causal forces is as it is" (17). This position grounds an extensive critique of Jonathan Haidt's social intuitionism and moral foundations theory, which I will discuss at length in Chapter Five.

between a moral life and one's individual flourishing, should one's community be such that this choice arises.

What is most relevant about Wong's project to an ethics for the depressed is his discussion of effective agency, which serves as the bridge between his criteria for adequate moralities and the role that an adequate morality plays in individual moral deliberation. Wong, reviewing Flanagan's critique of Williams, presents an account of possibility as a spectrum that would deemphasize the role of ground projects in identity formation and thereby complicate Williams's account of practical necessity. Instead of finding practical necessity internal to an individual, Wong locates it in a community's continuing contributions to identity formation. I conclude that the uptake of psychological realism by an ethics for the depressed should not involve centering integrity or a ground project, as Williams does, but instead accepting effective agency as a general constraint on adequate moralities and turning to the question of what the exercise of effective agency is like for a depressed individual and what, if anything, is missing. I do not aim to conclude once and for all if Wong is correct that there is such a thing as a general criterion for adequate moralities. But his discussion of effective agency is an illuminating guide for better understanding hypermoralized deliberation. In the case of depressed persons, I will go on to show, the central factor in hypermoralization is a diminished sense of agency.

2. **Wong on Psychological Realism and Effective Agency**

Wong (2006) expresses the core commitment of pluralistic relativism in a sentence: "There is a plurality of true moralities, but that plurality does not include all moralities" (XII). He limits the scope of adequate, or possibly true, moralities by reflection on "the functions of morality, human psychology, and the nature of human cooperation" to include those "requiring human beings to seek only that which they have some propensity to seek; inclusion of norms of reciprocity in light of strong self-interest; in specification of norms and reasons, balancing self- and other-concern in ways that include putting less pressure on other-concern through provision of some 'payoff' in terms of self-interest; justifiability of norms and reasons to the governed in terms of their interests when presented without falsification; and finally the value of accommodation of moral disagreement" (65).[36] Explicitly drawing on Flanagan, Wong affirms that his criteria for adequacy do little to predetermine the content of a true morality: "Morality is not *determined* by these deep human propensities, but if it is to serve as an effective guide to action, it must be limited by these propensities" (44). He later elaborates:

> universally valid criteria yield merely a skeleton of a morality, insufficiently rich in content to be action guiding. Consider the question of which priorities to place on … perfectionist versus egalitarian values; individual rights versus some form of group flourishing; individual rights versus utility; and obligations arising from special relationships versus impersonal obligations requiring, for example, respect

---

[36] I refrain from expanding on how Wong derives all his constraints on adequate moralities, via a mix of conceptual and developmental arguments about the points or purposes of morality, to focus on his discussion of effective agency. I note, however, that for a work in analytic metaethics, Wong is oddly silent on how he understands his constraints at a metaethical level, e.g., if he is a moral realist or thinks of these universal constraints as moral facts. (Paul Bloomfield's (2009) review in *Mind* makes a similar observation.) In any case, Wong's ideal theoretical strategy of deriving largely formal constraints on adequate moralities seems suited to his concern with moral cooperation and disagreements between and within communities, but less obviously relevant to an ethics for the depressed primarily concerned with an individual's moment-to-moment moral deliberation or internal conflict. I thus narrow my reading to the aspects of his work most relevant to my project.

for rights or promotion of utility. Within a moral tradition, there is room for a significant degree of indeterminacy … However, no society could afford to make these matters entirely "optional" in the sense of leaving to individuals the choice of what priorities to impose on each of these possible conflicts of values. The function of morality to facilitate and regulate social cooperation depends on a substantial coordination of expectations between individuals, and this coordination in turn depends on common expectations as to how others will decide to behave when important values conflict. (81)

Substantive rational disagreement over morals is possible on this pluralistic model of morality because the criteria for what constitutes an adequate morality leave much of the content of morality underdetermined. One's justifiable moral commitments may come into conflict with alternative commitments that are equally justifiable, both within and between communities.

Wong considers the objection of a moral universalist or absolutist that this very thought, that one's commitments may be equally as justified as rival (adequate) commitments, could sow doubt in the authority of one's own morality, yet responds that "we can be deeply committed to an ideal even while granting that not everyone needs to be so committed" (106). Contrasting switching moralities with switching the side of the road on which one drives, he draws directly on the themes of psychological realism, continuing: "To the extent that our lives take meaning and substance" from these commitments, "we will not be able to imagine ourselves having other kinds of commitments, even though we can perfectly well imagine and even observe that others have these kinds. There may be, with respect to many of our basic standards, no psychologically real question of whether we should adopt different standards or have none at all" (107). By invoking the "psychologically real," echoing a Jamesian conception of a "live option," Wong emphasizes the practical likelihood of conflict arising as a constraint on moral theorizing and the possibility – or necessity – of

185

ultimately justifying moral authority. Here the concept of psychological realism is deployed not as an attack on impartial morality but as a defense of partial morality against the likely concerns of impartial moralists. What matters for avoiding moral collapse, for Wong, is not the presence of an ultimate justification for morality but the absence of a threat to morality so serious that it could only be countered by such a justification.

But even though some form of psychological realism grounds Wong's defense against moral absolutism, Wong is circumspect about how to determine what counts as a "realistic possibility," the theory's central operative concept, when he turns to reflect on Flanagan's disagreement with Williams. In defense of Flanagan's more capacious and minimal standard of psychological realism, which undermines Williams's rebuke of impartial morality, Wong notes that "talk about realistic possibility is dangerous" in how it might "wed us to the cultural, political, social, and economic institutions that have shaped our current motivations" and "our vision of what is possible." But Wong also raises two cautionary notes against Flanagan. First, he argues, seemingly impersonal and demanding belief systems might not actually be held so stringently as they prescribe: Buddhism, for example, "imposes on its lay people rather modest requirements for action." The possibility of a personal commitment to impersonal values is weakened as a challenge to Williams, if it was one, by those seemingly impersonal values being, in practice, less demanding than they are presented as being. Second, some hypothetical possibilities, like raising the next generation to be morally superior to one's own, might be merely abstract: "teaching succeeding generations to be very different from us may amount to a mere abstract possibility, if it cannot be specified how that teaching would

actually make people different" (161). In introducing this distinction between abstract and realistic possibility, Wong positions himself as potentially more open than Williams but less than Flanagan regarding what sorts of possibilities might be considered psychologically realistic. Wong thinks that if psychological realism "suggests that moralities must take into account the content and strength of personal motives," as he does in dismissing the threat of equally justified alternative moralities, it must clarify how it limits "the *manner* in which people ought to be required to give impersonal consideration to the interests of others" and "the *extent* to which people ought to give such impersonal consideration" (159). The distinction between abstract and realistic possibility proves useful in providing this clarification.

Whether or not a possibility is "realistic" is a matter of degrees. Within the notion of a realistic possibility, Wong further distinguishes between a conceptual aspect, "a conception of the *process* by which it [the possibility] could be realized," and an empirical aspect, "*evidence* for the relevant agents' ability to initiate and complete that process" (162). In turn, a possibility might be unrealistic either because we do not know how it would be realized or because, even if we do know how, there is no reason to think, or good reason not to think, that it would be realized. There are degrees of clarity both about the process and of the evidence for the process being realized under specific circumstances; moreover, those circumstances can change, and with them the availability and meaning of the evidence. If we are to ask if commitment to an impersonal value is psychologically realistic, Wong thinks, this is partly a question of how "to mediate the commitment to the impersonal through sufficiently strong personal commitments," that is, "the *manner* in which that requirement is to be fulfilled" (170). It follows for Wong

187

that the success of impersonal values depends on "institutions and practices that embody and facilitate on the individual level the mediation of impersonal values through personal values" (175-176). There may be lofty moral ideals that are "real enough" to be worth striving for, such that a commitment to psychological realism does not disqualify these ideals from being prescribed, but demanding enough that individuals should not be blamed for failing to live up to them without a morally motivating support network. Wong's version of psychological realism thus has, I think, a firm position regarding the manner of moral commitments: impersonal values or ideals, to be authoritative or bindingly obligatory, must be compatible with some sort of mediation through personal motives or goals. If this means that moral theories should not require indifference to the presence or absence of personal motivation, then this preserves the foundation of Williams's challenge to impartial morality.

But as to whether psychological realism restricts the *extent* to which impersonal values can require overriding more personal values – if, for example, an adequate morality can require us to sacrifice our (other) treasured projects and attachments or our integrity – Wong finds the theory "inconclusive" (160). First, the contingency of possibility complicates any attempt to demonstrate that a moral theory asks too much of agents. What is too much to ask for us here and now might not be too much to ask generally, or elsewhere, or in the future, or in the past. Second, it is not obvious how a psychological realist should respond to indeterminate evidence of the possibility of a moral requirement, and much evidence is likely to be indeterminate. If a very demanding moral requirement is, say, "probably not" possible to personalize and pursue on the best evidence available, yet it is not necessarily "totally unrealistic," a psychological realist

could self-consistently recommend moderating the requirement, or pursuing it precisely to push the bounds of what seems possible, or rejecting it to guard against alienation arising from repeated moral failure.

Put this way, detached from the centrality of a ground project in Williams's version, psychological realism has something to say about the form of morality – and something significant in the challenge it poses to impartial moralities – but little to say about its content. The "constraints of realistic possibility" require "cultivating effective moral agency" (175), which in turn requires appropriate institutions and practices, and realistic possibility is a matter of degrees, the appropriate response to which is underdetermined. The one general constraint on adequate moralities that does emerge is that "moralities that in some way depend for their acceptance on denying the reality of certain possibilities" that are in fact "*real enough* (if not realistic)" should "be ruled out as inadequate" (176). In a twist, the payoff of psychological realism, rather than being that moralities should not demand too much, is that they should not underestimate our capacities. Beyond this, Wong's version of psychological realism, which restricts the manner but not the extent of moral commitments, becomes – though he does not describe his conclusion in these terms – coextensive with his theory of "effective" moral agency.

Psychological realism amounts to effective agency for Wong not just because what counts as a realistic possibility for us is, functionally, what choices or options we can think of and effectively act on, but because effective agency is the core of Wong's argument that "there must be a personal perspective" because of "our social nature as human beings" and there is "a universal constraint on adequate moralities that they must contain such a perspective and its attendant duties." Contrary to an impersonal or

189

impartial moral theory that would claim that moral authority requires a general overridingness of moral obligations over sentiments or attachments, severing moral goodness from prudential interests, he defends a basic commitment of a personal approach to ethics, that "the reason for performing some of our most important ethical duties lies in our own identities and in our flourishing" (116).

Wong's concept of effective agency is meant to show how "our social nature" grounds at least some of our reasons for fulfilling our moral duties. He defines effective agency as "the set of abilities that allow us to formulate reasonably clear priorities among our ends, and to plan and perform actions that have a reasonable chance of realizing our ends, given all the conditions beyond our control." Effective agency has a social aspect because the "necessary conditions for effective agency include the possession of certain relationships with others that in turn are partially defined and sustained by duties we have toward them." In sum, "flourishing requires effective agency," which requires certain relationships, which require certain moral duties (119). Wong thinks that while this argument links some moral duties to flourishing, relying on the necessity of certain relationships to flourishing, it is compatible with diverse and competing understandings of flourishing and which, or how, moral duties follow from it.

Drawing on his work with Amelie Rorty, Wong argues that the relationships necessary for effective agency lead agents to possess a "practical identity," a collection of traits that are "central to identity" with no necessary connection between them but generally sharing at least some common features (120). Traits become central to one's identity, and so feature in one's practical identity, in many ways: the extent to which other traits or dispositions depend on it; to which the trait manifests in different contexts

and relationships; to which the trait is difficult to change; to which the trait affects how agent is treated or categorized by others; to which it affects the agent's resilience; to which it supersedes other traits in internal conflicts; or to which the agent sees themselves as changed when or if the trait is changed or lost. Traits that feature in one's practical identity might also be relevant to an agent's self-evaluation or be actively developed or strengthened by the agent. Practical identity traits also affect how an agent may act in multiple respects: they may affect what seems or is perceived as salient to an agent, forming "the problematic of their experience"; they "propel agents into certain sorts of situations and problems," affecting "the dynamics of social interaction" by how others respond to them; they may affect or explain an agent's acquiring certain beliefs, patterns of desire, values, or goals requiring deliberation; if they influence child-rearing or socialization, "they direct the formation of habits" (124). From a theoretical perspective, practical identity serves as the evidence of how an agent's relationships – their affiliations, their roles, their loyalties – shape and structure their effective agency.

If the core of Williams's version of psychological realism, the consistent contingency that takes the place of necessity and acquires its authority-anchoring and deliberation-structuring functions, is a ground project, the core of Wong's version is relationships and their "necessity." A practical identity is defined by social relationships that may or may not be conducive to flourishing, as when one's ideal identity clashes with the identity ascribed by others, so not all practical identities or the relationships underlying them facilitate effective agency. But the practical identities that do, which Wong calls "effective identities" (128), involve "adequate knowledge of social norms, reasonable congruence" between different aspects of one's identity (e.g., ideal and

191

social), and "self-esteem" (129). Insofar as acquiring these three qualities requires nurturing and care, "the nurturing of an effective identity in an individual requires relationships with others that are partly defined and sustained by reciprocal duties." The "powerful hold" of "special duties to particular others," like parents toward children and children toward parents, arises because "moral agency must grow within the context of relationships governed by such duties" (138). Moral agency "is a species of effective agency" for Wong, the "ability to formulate reasonably clear priorities among one's moral ends, and to plan and carry out courses of action that have a reasonable chance of realizing those ends." Insofar as "the function of all moralities" is "the facilitation of social cooperation and the guidance of individuals toward worthwhile lives," they thus must "promote relationships that involve identifications, trust, and reciprocal care" (143). The reasonable congruence between different traits and commitments required by an effective identity implies, insofar as moral agency is a subspecies of effective agency and "morality requires effective agency" (144), that traits typically considered moral must be weighed against traits typically considered prudential in order for all traits involved to retain their motivational efficacy. Thus, the concept of effective agency grounds a personal perspective toward morality, rejecting any general overridingness of moral over ostensibly non-moral obligations.

But Wong's account of relationships as "necessary conditions" for effective agency is a bit slippery, because what might be necessary for the development of an individual, e.g., from childhood to adulthood, might not be necessary for their persistence as an individual and everyday exercising of agency. Wong's concern is largely at the level of the social: he wants to account for how different communities, and

192

subcommunities within communities, can have divergent and equally legitimate moralities and so rationally disagree with one another. The threat of the individual amoralist who, upon developing effective agency, turns their back on their community and refuses to participate in pro-social activities simply, in the account offered thus far, does not arise. Wong's iteration of psychological realism seems the most fully developed of those available. But in trying to adapt Wong's version of psychological realism to an ethics for the depressed, the question of how exactly relationships are "necessary" for effective agency and motivational efficacy becomes sharper. Specifically, the distinction mentioned above, which Wong does not emphasize, is important: someone might be raised properly, in the sense of having the relationships necessary to develop effective agency, but then become a less "effective" agent by becoming depressed or just alienated from society. If this depressed or alienated person, seeking justifications for going on in the hope that this exchange of reasons will spur motivation within themselves, turns to Wong for ethical arguments about what rules bind them and why, what, by the lights of pluralistic relativism, should he say? Even granting that the issue of moral bindingness or authority is more a problem for a community's practices of moral development than a theoretical conundrum, the question of what can be justifiably said in defense of norms to an alienated individual remains open and significant.

Impressively – and, to me, unexpectedly – Wong later turns to this question and makes a striking claim: while it is possible to justify to an alienated individual that a moral life is compatible with flourishing, it is not possible to justify the claim that flourishing requires a moral life, meaning that pluralistic relativists cannot argue for morality as a necessity condition of a well-lived life. Wong acknowledges that his theory

193

"will remain disturbing to those who insist that our justifications for morality must engage the motivational structure of each individual" and that he does not promise "an answer to the sociopath who asks the question 'Why be moral?'" (200). He then digs deeper into the question of moral authority for a pluralistic relativist: "if morality does not constitute an irreducible part of the world's fabric that simultaneously has prescriptive authority over us, why go on accepting the influence it has over us?" Wong thinks that rejecting all moral obligations is generally not a live option. But he recognizes that this does not settle the matter: not only may some become alienated, but one might come to see the inescapability of a morality's influence or call as bad, perhaps harmful to the prospects of a well-lived life. It may even be characteristic of morality to be experienced as an imposition: "we know that morality can call upon us, and that in its name we can call upon others, to make great sacrifice" and perhaps we "need to see that sacrifice as required by something greater than ourselves" (202). In the absence of another objective foundation for morality, a strong eudaemonic strategy, arguing that one cannot flourish or truly be fulfilled if one is not moral, is tempting. But Wong firmly rejects this, presenting as counter evidence

> the widespread conviction that one can insulate oneself from the problems of the larger society. A just world would be one in which one could not succeed in this endeavor, but I see no evidence that the actual one is such a world. Closing the gates can work if one has enough money and if a moderate degree of luck holds. I have seen no attempt to show that individual flourishing requires morality that is immune to these kinds of problems … induction is enough to move me to a more modest project: showing that there are forms of flourishing that involve a moral life, even if not all forms of flourishing do. (205)

For Wong, it simply is possible to be immoral and fulfilled. The arguments for social and moral norms that one can make to an alienated individual are reasonably constrained by this fact.

194

But if the criteria for what constitute adequate moralities emerge at least in part from developmental claims about what is conducive to flourishing, then it is plausible that at least some adequate moralities should be compatible with flourishing. It would be consistent with Wong's pluralistic relativism to respond to an alienated person with something like: "There is probably a community out there with social and moral codes of conduct that I could respect and that would better meet your needs as an individual." It is this modesty or moderation of Wong's theory, its centering of rational moral disagreement as the fact around which morality should be understood, that makes it so relevant to an ethics for the depressed. As I have argued, depressed persons often seek justifications when their problem is motivational and ruminate more than deliberate, greatly limiting what an ethical theory can do to help them. But adopting a way of thinking about ethics that loosens its bindings could be, if anything is, a way of alleviating existential guilt.

One major question that remains is how Wong's core concept of effective agency should be taken up by an ethics for the depressed. By the standards of effective agency, which depends on one's practical identity and so incorporates many contextual and environmental factors in its definition of "agency," it seems fair to say that a depressed person is a less "effective" agent. But if my arguments about depressive loss of motivation are correct, then the question of if a depressed person "loses" moral agency is at least complicated: an agent might contend with a compulsive habit or quasi-belief that frustrates their actions and makes them a less "effective" agent without, in another sense, being any less rational or an agent. I will leave evaluating the concept of effective agency as a definition of agency to others. Instead, I will try to pinpoint what specifically is

195

happening when a depressed person "loses" agency, and how that might affect what ways of ethical thinking, if any, could ease or prevent this seeming loss.

3. Gallagher on the Sense of Agency

Shaun Gallagher's (2020) enactivist theory of action in *Action and Interaction* synthesizes two decades of research to present a close examination of the phenomenology of agency, including the topic most relevant to experiences of depressive rumination: the sense of agency not just in action, but in intention-formation. I set aside any discussion of the merits of enactivism or how this account of agency as experienced supports that theory to focus on Gallagher's fine-grained account of the sense of agency itself. I deploy this account to argue that the sense of agency serves as a little-noticed epistemic guide in deliberation, offering an affirmative intuition that one is "thinking for oneself" when forming intentions, and this guiding intuition is weakened or minimized in experiences of depression or anxiety, making it difficult to determine when to conclude deliberation.

Having a sense of agency is not necessary to know, believe, or warrant a belief that one is an agent. But, practically, it is difficult to confirm deliberatively that a course of action is in line with one's goals or values if one lacks the felt and implicit verification that they are acting "for themselves" and, say, feels instead as though someone else is acting through them or that "no one in particular" is acting at all. I think this is what is happening moment-to-moment in depressive rumination that fails as deliberation by failing to lead to action. The same pre-intentional states, quasi-beliefs or habits of feeling, that inhibit motivation also inhibit a depressed person's sense of agency when they are forming intentions, leading them to feel alienated not just from their long-term goals but

196

even from short-term intentions as they are being formed and developed through deliberation. Moreover, I argue, maintaining in practice a distinction between the moral and the (merely) prudential involves an awareness of "my interests" or "my obligations" usually implicit when reflecting on and resolving conflicts between them. One may understand this distinction abstractly but struggle and suffer to apply that understanding if one cannot rely on it feeling obvious that one's intentions are one's own. The temporal dislocation of existential guilt, the perceived impossibility of a redeemed future, leads to this self-alienation. The challenge a depressed person faces with "my obligations" possessing a special moral status is less to do with the "obligations" and more with the "my."

Gallagher claims that action is "characterized by three elements or aspects" beyond bodily movement: "intention, a sense of agency, and meaning which generally goes beyond the agent's intention" (43). The sense of agency can be further distinguished from the sense of ownership, both phenomenologically and neurologically. In involuntary movement, as in the experience of being pushed by someone else, one has a sense of ownership but not a sense of agency, resulting in an intuition like "I am the one moving" but "I did not cause the initial movement or have immediate control of it" (44). These motor-control level senses of agency are phenomenologically recessive or pre-reflective, intuitions arising without need for special attention or introspective reflection, and they are ongoing, arising "not simply when I initiate an action" but perpetuated by "continuing efferent signals, and the afferent feedback that I get from my environment" (45). But insofar as one is aware of one's own action, one is typically aware of a more abstracted intentional aspect, that being "what" one is doing, like writing a sentence, and less so

197

"how" one is doing it at a motor level, like moving one's fingers across a keyboard. Thus, understanding what it means to be aware of forming an intention is complicated by the different kinds of intentions there are at different levels of abstraction.

Gallagher distinguishes between three kinds of intention: distal or D-intentions, proximate or present P-intentions, and motor or M-intentions. Distal intentions are longer-term goals or plans prospectively formed through deliberation. Proximate or present intentions are intentions "in" actions, or reflectively or perceptually monitored or guided actions understood in terms of means-ends relations. Motor intentions are processes that control movements that carry out intended actions, typically without conscious perceptual monitoring. Given the descending levels of abstraction and awareness in this taxonomy of intention, a further distinction becomes key to the phenomenology of agency: second-order prospective reflection, which may involve evaluative judgments, is distinct from first-order pre-reflective experience, where a sense of agency may be felt or intuited. One may judge oneself to be an agent by taking one's sense of agency to be supporting evidence, but "the sense of agency is not itself a judgment" (49). Even so, the sense of agency may be strengthened through the formation of D- and P-intentions: an agent might reflectively craft a narrative about their own decisions which then intensifies their felt sense of being an agent. Gallagher refers to the judgment that one is or has been an agent as "retrospective attribution" and notes that while, again, the sense of agency is not constituted by conscious attribution, one might experience a minimal sense of agency "because there is no D- or P-intention and no retrospective evaluation," merely "my thin experience of having motor control over something I am doing" (51). In an earlier paper, Gallagher (2013) cites narcotic addiction

198

as an example of such a disruption: "If a drug addict invests himself in resisting drugs he may feel that something other than himself is compelling him to drug use. If he withdraws from taking the drug, when he starts using again he may not conceive of himself as the agent" (12). The interference of compulsions, by disrupting the formation of D-intentions and deliberation, can undermine the felt sense of agency as well.

Gallagher and Trigg (2016) connect this line of reasoning to agoraphobic anxiety, emphasizing that a diminished sense of agency (SA) or sense of ownership (SO) can lead to doubt in one's agency, as a loss of evidence to support the judgment that one is acting as an agent, and so actually diminish agency. The authors present a bodily inhibition model of anxiety, observing that in agoraphobic anxiety, "a disruption in agency can lead to a disruption in a sense of self more broadly" in two ways (6). First, there is often a disturbance in "bodily motricity," this referring to Merleau-Ponty's term for "the body's action-oriented power to project intention into the world in a movement of spontaneity and possibility" and which the authors regard as "the general source of SA" (8). This seems to map on to what Gallagher refers to as M-intentions in his 2020 book. In such disturbances, "sensations of anxiety, including the inability to move or the sudden urge to move, is felt as if it comes from nowhere" (6). Second, one's own body is encountered as a thing or foreign object rather than lived through as a center of subjectivity, as *korpor* rather than *leib*, thereby partially disturbing the sense of ownership. The "partial (but never absolute) loss of SO" can involve parts of one's body seeming to take on a life of their own: "it is not that I am the one running from danger, but rather it is the legs that are instructing me to run" (9). Generally, if one is not afflicted with chronic anxiety, "these movements of self-alienation and bodily objectification are brief, and are often

199

consolidated into a unified and relatively coherent sense of self that includes SA and SO." Agoraphobic anxiety, manifesting as bodily inhibition, leads instead to compensatory habits of reliance on "familiar objects" or "means of escape," while "the inevitable failure to maintain this tightly woven yet precarious grip on control leads to anxiety" and "a diminished SA" and "partial disruption in SO" (8). This amounts to a vicious circle: anxiety disrupts self-identification, one's body seems not fully one's own, and one develops compensatory habits that attempt to ward off anxiety but deepen one's senses of powerlessness and alienation when one's practices inevitably fall short of their goal.

Gallagher's (2012) reflections on depression and the experience of time, consistent with Ratcliffe's (2015) claims in *Experiences of Depression*, help to link this account of agoraphobic anxiety to the matter of depression: depressive disruption of self-identification most plausibly emerges not from a partial disturbance of SO, as in agoraphobic anxiety, but of temporal experience. Gallagher observes that depression intensifies short-term experiences of sadness, seeming to expand a depressed person's emotional range, while flattening longer-term experiences of sadness like tragic films, seeming to restrict that range. He infers that because depressed persons have a "slowed experience of time flow" and "tend to overestimate time spans," being "preoccupied with past events" and "less focused on present and future events," whether depression adds or subtracts depth from experience depends on the temporal experience involved (128). If a depressed person must be able to, say, appreciate the loss of a possible future to be moved by a tragedy, and this capacity for futural anticipation is deadened (especially for themselves, but perhaps also somewhat generally), a flattening rather than deepening of

sadness makes sense. Furthermore, if an undisrupted identification with one's own goals involves prospectively identifying with one's future self who has accomplished said goals, a disturbance of temporal experience could inhibit the generation and development of D- and P-intentions.

It would thus be a mistake to think of the sense of agency as an epiphenomenon, a curious offshoot of the real philosophical matter of what agency is, especially in developing an ethics for the depressed. True, a definition of agency will not follow straightforwardly from what agency feels like. But there are, generally, bidirectional causal relations between the sense of agency and agency itself, or at least effective agency. The presence or absence of D- and P-intentions can bolster or diminish the sense of agency, and a bolstered or diminished sense of agency can affect how an agent retroactively attributes agency to themselves, producing virtuous or vicious cycles that may actually lead to a broadened or restricted capacity for action. More to the point, the problem that urgently confronts a depressed person is not clearly defining agency but how to muster enough confidence in their own agency, in a broad and underdetermined sense, to go on. Understanding their dilemma, which pertains to the sense of agency and its structural conditions, does not require grounding this confidence in a correct assessment of one's agency or knowing the criteria for said assessment. If we accept that the sense of agency is a central epistemic guide or obstacle in deliberation, then the sense itself is worthy of more philosophical attention. As is, the inattention to the phenomenon, Gallagher excepted, is striking.

Gallagher's (2020) research into the sense of agency in intention-formation reveals a central ambiguity in the sense of agency generally: how we experience agency

is relative to how we individuate specific actions. A process or element of deliberation can lack a sense of agency, as in the case of experiencing a thought as impulsive or unbidden or, more seriously, a schizophrenic experiencing a thought as inserted. This may have a downstream effect on one's judgments or retrospective attributions of one's own agency as one's self-concept is undermined by the evidence of one's disturbed intuitions. But it is also the norm for there not to be a sense of agency for every process of action monitoring or deliberation. Gallagher offers the example of "putting together a piece of furniture by following a set of instructions":

> I could have a sense of agency for following the instructions and closely monitoring my actions in terms of means-ends. Certainly doing it that way would feel very different from doing it without following the set of instructions. But the sense of agency for following the instructions would really go hand in glove with the sense of agency for the action of assembling the furniture. How we distinguish such things would really depend on how fine-grained we get in defining the action. (53)

He thus cautions against erroneously reading "legitimate conceptual distinctions" back into the phenomenology, that is, developing a too fine-grained taxonomy of distinct senses of agency when, in practice, only the sense of agency and sense of ownership are readily distinguishable (55). There might be, for example, some difference between a sense of mental effort and a sense of physical effort, but it is also plausible such hair-splitting is the projection of theoretical categories into a more ambiguous first-order phenomenology. He continues that, conversely, the phenomenology should not be oversimplified: for example, there is no such thing as a "naked intention," presenting itself as felt by no one at all, as even cases of "schizophrenic delusions of control" involve the subject "attribut[ing] agency to someone else" (56). Gallagher advises the careful drawing of methodological lines between phenomenological analysis of what is

given in experience, neuroscientific analysis at the neuronal level, and conceptual analysis of cognitive models or psychological motivations and narratives (58-59). In so doing, the distinct contributories to what manifests as a sense of agency can be discerned.

Gallagher concludes that there are at least five such contributories to the sense of agency associated with an action: the formation of D-intentions, or prospective deliberation; the formation of P-intentions, or conscious goal-oriented action monitoring; basic M-intentions linked to bodily motility; perceptual monitoring of the effects of action; and the judgment or retrospective attribution of agency following the action. The sense of agency and sense of ownership typically do not come apart but can and, in so doing, produce distinct intuitive feedback, as in the case of the experience of involuntary movement. A sixth contributory may be one's "long-term sense of one's capacity for action over time, related to self-narrative" (59).

Following Gallagher's research, I propose that a depressed person's diminished sense of agency primarily involves a disruption of temporal experience that undermines D-intention formation: deliberations become difficult to conclude without a clear intuited sense that one's future self who accomplishes an action is the same self who is deciding to act now. This loss of sense of agency results in a lack of epistemic guidance that slows and complicates decision-making, though does not render it impossible, as one's established and enduring beliefs in oneself as an agent may overcome one's self-doubting quasi-beliefs or feelings. But over time, a diminished sense of agency can, as intuited evidence on which the depressed person is likely to ruminate, diminish those beliefs in oneself as an agent as well. The problem turns on the capacity to maintain a teleological throughline, at the level of felt or intuited identification, of what "my purposes" and "my

interests" are. Existential guilt is depersonalizing, detaching one's perceived

unworthiness from any specific act or trait, and thus any deliberation regarding if "my

obligations" override "my interests" becomes confounding just because it becomes a

painful struggle to clarify "my interests" over time. The most appealing recourse is to

throw oneself into altruistic impulses, however well-realized or effective, at the cost of

self-neglect. The result is hypermoralized deliberation, in which all prospects appear as

more or less adequately altruistic.

Guilt seems to withdraw the future, disrupting time. The disruption of time

confounds future self-identification, which undermines prospective deliberation. And if

moral deliberation is defined in terms of weighing moral obligations and prudential

interests and successfully valuing the former over the latter, depressed persons will fail at

the weighing, though not the valuing. The depressive collapse of desires into obligations,

of "want" into "must," and the shroud of guilt coming to blanket every affordance does

not just follow automatically from the nature of that guilt: it occurs as the depressed

person's ability to think clearly about what is prudential for them, as a person with

meaningful future possibilities, erodes and leaves impulsive altruism entwined with

despair. The weighing of moral deliberation will likely not be rationally concluded but

either languish in rumination or be desperately dodged through some moral-seeming self-

abnegating act. When a strict theoretical moral-prudential distinction is put into practice

by a depressed person, if I am right that this same person is likely to take it as a decision

procedure too, then it is likely to intensify all their most self-devastating patterns of

thought.

If it is at all the business of ethics to provide a model for how a depressed person might think differently, then moral obligation should be reconceived without a strict hierarchy of moral over prudential. I conclude this chapter with some reflections on what I call "ethics as a reliable guide," or what Samuel Scheffler refers to as a "moderate" approach to ethics, and alternative approaches to thinking about moral authority and its attendant moral-prudential distinction.

4. Ethics as a Reliable Guide

Depressed persons suffer unduly in their experiences of moral obligation. Some of this suffering is indifferent to ethical theory, to the way in which "moral obligation" is understood, because a depressed person is a compulsive moralizer and is not likely distinguishing between moral and non-moral obligations as any theory would prescribe. For them, contradictions arise in deliberation between rational judgments about what prospective actions are possible and quasi-beliefs or compulsive thoughts of those actions' impossibility. When this generalized sense of the impossibility of success is tied to one's perceived personal unworthiness, a dreadful sense of urgency arises: this unworthiness presents itself as the reason why success is impossible, and so if that impossible success were somehow achieved, an impossible redemption from one's unworthiness would follow, a sign that one must have been worthy all along after all. The stakes of the smallest task thus come to seem immensely great. Basic desires or interests take on the appearance of heavy duties. The capacity to go on comes to appear from within as the right to go on, the elusive evidence that one is more than a broken husk.

It is obvious that a person in this frame of mind has restricted psychological capacities. One might think that morality itself, especially one framed as impersonal and impartial, could be asking too much of someone so compulsively ridden with guilt. As the literature on psychological realism shows, there are good reasons to reject an impartial view of morality: the attempt to justify moral authority by eliminating moral luck, while well-intentioned, may be both alienating and futile. If so, the theoretical positions involved in eliminating moral luck – that there is a special kind of moral obligation, that it generally overrides prudential interests, that there are purely voluntary acts, that there can be perfectly just evaluation of purely voluntary acts – may also fall away. But the case of the depressed person does not straightforwardly provide more reasons to reject an impartial view of morality. If ethical thinking plays a role in why they may be alienated or may have lost integrity, it is probably because of their compulsively perceived, not their actual, obligations. Their undue suffering does not arise because morality is excessively burdensome to their personal interests or projects. Rather, they throw themselves into morality to the exclusion of their own interests or projects, finding in their altruistic impulses a motivation that can overpower, however fitfully, their malaise.

The reason depressed persons suffer in their experience of moral obligation, then, is not because of the burdens of the moral but because of the felt insubstantiality of the prudential. In depressive deliberation or rumination, the prudential dissolves into the moral, every desire and interest becoming like a shard of a mirror reflecting one's own moral fallenness, powerlessness to succeed, and impotence as an agent. But a coherent category of the prudential must be sustainable in practice for the moral to have a special

status overriding the prudential. Depressive moral deliberation, with its subversion of prudential concerns, cannot rationally proceed or conclude in that manner, and so spins its wheels. If an ethical theory can adapt to these conditions and provide a more functional model for deliberation, it must explain how deliberation about one's obligations can do without a hierarchy of the moral over the prudential, adapting to a context in which the prudential is unthinkable.

There are precedents for this: Julia Annas (1995), for example, proposes a reinterpretation of ancient philosophies as lacking a modern notion of the prudential as conflicting with the moral. Whereas "in many modern theories it is taken for granted that the agent will, as a matter of prudence, aim at his own self-regarding good," Annas claims, "in ancient theories rational reflection on one's life as a whole leads the agent to reason morally, and no distinct competing role is left, within the theory, for prudential reasoning" (242). She rejects the interpretation that in such theories moral reasoning is just an extension or development of prudential reasoning, like, say, an individual continuing to pursue their own interests but having a more enlightened understanding of what those interests are. In the case of Stoicism, which she takes to be generalizable to other ancient eudaemonist theories, she argues that there is a "distinctive change of perspective" that occurs in ancient ethics, a "recognition that reasoning has its own characteristic aim, which the agent values for itself, and is not just valued for the results it assures for the agent" (251). There is a meaningful shift between what we moderns might understand as prudential and moral thinking, but there is no sense of an agent undergoing a perpetual conflict: either one is unenlightened and thinks in terms of petty self-interest or one has become enlightened and grasped their true calling as a rational being. Annas's

reading retains a theoretical hierarchy of moral over prudential rather than dissolving the distinction. But both Annas's interpretation and the one she rejects present ancient eudaemonic ethics as self-transformative and as lacking a practical (that is, applied in deliberation over moral dilemmas) hierarchy of moral over prudential considerations.

One path for an ethics for the depressed is thus to develop a mode of philosophy that is transformative or therapeutic. These are two modes of "philosophy as a way of life," to borrow Pierre Hadot's phrase, distinct from what I would call orthodox or argumentative philosophy that aims just to change beliefs through rational argument. I understand a transformative ethics as one that aims to change the dispositions of agents to foster or ease ethical action, and a therapeutic ethics as one that may not aim to change an agent's dispositions but aims to sustain the agent by allaying the social, emotional, or cognitive burden of ethical behavior. The most obvious virtue of orthodox philosophy is that its success or failure is not, or at least is not supposed to be, contingent on the emotional state of the contemplator. Transformative and therapeutic ethics instead involve, as Hadot has called them, spiritual or existential practices. If they are not transformative or therapeutic, they presumably fail on their own terms, at least for a given agent.

I want to see how far orthodox philosophy can go in constructing an ethics for the depressed. I think that depressed persons seek reasons to go on in the form of arguments, and while those reasons may dash against the rocks of implacable compulsion, giving a good account of that process itself involves making arguments, as I have here. Before attempting anything like transformative or therapeutic ethics, I want to present the reasons that support how I frame depressive loss of motivation or disrupted deliberation,

208

both to subject them to rigorous analysis and, if they are rigorous, to aid in the self-understanding of those who find the framings useful. My approach to argument is one that assumes a therapeutic goal on the part of the reader, but which responds by taking argumentative rigor as seriously as possible. I think that orthodox philosophy might benefit, become kinder and more generous, if more philosophers shared this assumption. But as for transformative ethics, at the risk of being glib, I fear that evoking transformation through theory is an endeavor for sages. The aspiration to sagacity is noble, but here it would be, let's say, premature.

Outside of emulating ancient eudaemonic ethics, then, a good candidate for a modern reframing of the morality-prudence distinction is Samuel Scheffler's (1992) account of stringent and moderate moralities and their guiding Ideals of Purity and Humanity in *Human Morality*. Scheffler defends a moderate conception of morality, arguing that "although moral considerations and considerations of self-interest can diverge, and although morality sometimes requires significant sacrifices of us, nevertheless the most demanding moral theories are mistaken." A moderate morality is rooted in an Ideal of Humanity "according to which morality is, from the standpoint of the individual agent, fundamentally a reasonable and humane phenomenon." A stringent morality is by contrast rooted in an Ideal of Purity, for which "morality's concerns are highly distinctive and sharply opposed to the self-interested concerns of the individual" (6). The appeal of the stringent view of morality derives, Scheffler thinks, from two factors: its sharp distinction between morality and prudence and its ideal possibility of self-transcendence. To see morality as fundamentally difficult and demanding, or moral agents as fallen or otherwise chafing under morality's constraints, is also to affirm that

those who transcend their merely prudential interests have transcended their impure nature. Scheffler acknowledges the appeal of this view while arguing in line with a moderate position that living morally is a realistic possibility and, moreover, that it is better to see morality as such, so as to affirm the value of personal goals and attachments and the possibility of integrating them with higher self-transcending ideals in a coherent moral life. The hierarchical conception of moral considerations as overriding prudential interests is reframed as a component of a stringent morality, echoing Williams's assessment of impartial morality or the morality system.

Scheffler takes on the subject of overridingness directly, and rather than rejecting the notion, deftly renders it relatively toothless: morality may be considered overriding so long as it is moderate, congruous with a well-lived human life, and thus not excessively demanding. If belief in the overridingness of moral obligations may be summarized as the view that "it can never be rational knowingly to do what morality forbids" (7), then Scheffler thinks that "just because it is so strong," the claim is "unlikely to be true" (56). Even so, he is not convinced by the usual arguments against overridingness, which he summarizes as the objections to the view that "morality requires humanly incorrect behavior" (58). He thinks that objections to overridingness are rather objections to stringency, that is, to models of morality like Williams's morality system. If morality is moderate, and moral obligations are conceptualized as already limited by and accessible to human capacities, then the notion of moral obligations being overriding ceases to be objectionable to these objectors. The problem is not with an imperative to consistently value moral considerations over, say, self-interested whims, but with a view of morality that regards all non-moral considerations as little better than self-interested whims when

weighed against moral obligations. Similarly, Scheffler argues that the pervasiveness of morality, the notion that any and every aspect of one's life is in principle a possible object of moral evaluation, is not objectionable as long as that evaluation is conducted in the context of a moderate ethical system, which is already inclined to view morality as less strict. Whether the metaethical particulars of an ethical theory, like if it is overriding or pervasive, are demanding or innocuous depends entirely on if the content of that theory qualifies as stringent or moderate.

Wong (2006) critiques Scheffler's discussion of the Ideals of Purity and Humanity, but he is ultimately arguing for a kind of moderate approach to ethics himself. Scheffler supports a moderate over a stringent approach by arguing that the "the Ideal of Humanity can accommodate to some extent the Ideal of Purity, but not the other way around," the latter being considered "a supererogatory ideal" (26). Wong points out that this understates the tension between the ideals: if, for example, the Ideal of Humanity supports the claim that special duties, e.g., to parents or children, can override all other duties, an impartial and self-transcendent commitment to treat all equally in line with the Ideal of Purity would constitute a rejection of the former ideal, not an especially high aspiration consistent with it. Wong concludes that there is "no determinate ranking within our tradition that orders these two values" and "no compelling, deeper rationale for subordinating one to the other" (27). Wong, like Flanagan, is reluctant to delimit the content of what counts as an adequate morality, focusing on formal requirements. But as I have pointed out in the discussion of psychological realism, this can put Wong and Flanagan in the peculiar position of endorsing much of the content of a belief system while rejecting that belief system's self-conception. If an adequate philosophical ethics

claims to be derived from axioms of pure reason and that alternative theories are delegitimized by rational errors, or an adequate religious ethics is justified by claims to the religion's exclusive divine truth, a pluralist must affirm the ethics while rejecting the justification offered for that ethics. This is self-consistent. But it shows that a commitment to pluralism, to a kind of metaethical moderation, also entails rejecting some stringent claims even when they are part of the content of an ethical system.

This is "ethics as a reliable guide": prioritizing self-persistence over self-transcendence, consistency over necessity, the provisional over the certain, and more streams of motivation over better-justified authority. The acceptance of some level of moral luck implies the imperfect apportioning of credit and blame, but as Williams and Wong argue, a humble commitment to be as fair as possible is compatible with the impossibility of perfect justice. If a moderate or reliable ethical theory is one that may do without a hierarchy of moral over prudential considerations, and if a theory qualifies as moderate or reliable based on the content of its prescriptions and proscriptions, then the next task is to be as clear as possible about what the content of an ethics for the depressed should be and why. In the next chapter, I will argue that a reliable ethics for the depressed, for those struggling with demoralization and hypermoralized deliberation, is a value ethics that prioritizes intuition over procedure and views the latter as a development of the former.

CHAPTER V: A VALUE ETHICS OF ENGAGEMENT

A depressed person has an ambivalent relation to their own ethical intuitions. Depressed persons are highly motivated by altruistic concerns, which may be experienced along with a strong intuition of having a personal responsibility to act in accord with these concerns. But this motivation may be inhibited in its development into intentions or actions by compulsive quasi-beliefs or habits of feeling, a process I have called demoralization. This may lead a depressed person to see themselves as unmotivated even as they direct their motivation into rumination or spiraling, which I have identified as a hypermoralized deliberation in which the distinction between the moral and the prudential is effaced as a depressed person devalues their own projects or purposes alongside a dampening of the sense of agency involved in intention formation. One set of ethical intuitions that esteem other persons is undercut by another set of ethical (or guilt-adjacent, and so ethical-feeling) intuitions that denigrate oneself. These sets are separable in theory, but in existential guilt they are experienced as a whole or circuit, though the aspects that seem dominant or driving will be ambiguous or vary. For example, I may feel that I must aid others because I am unworthy, or I may feel that I am unworthy because I have failed to sufficiently aid others. The success of an ethics for the depressed depends, I think, on if these sets of intuitions are separable in practice and to what extent that practice may be theorized.

An impartial ethics may appeal to a depressed person's impulse to be strict on themselves in the hope that ethical action can overcome existential guilt. I have tried to show how this impulse involves a category mistake born of compulsion and desperation, one that conflates motivation and justification by responding to a loss of motivation with

213

a search for justifications and treats existential guilt as though it were intentional guilt capable of being alleviated through an act of forgiveness. The theoretical challenge of eliminating moral luck and achieving a truly impartial morality system represents, in the context of an ethics for the depressed, the existential challenge of knowing what measure of one's own ethical worth one can trust. To this existential challenge, the response must be largely indirect, that is, therapeutic or exhortative: the answer is not to refute one's intuitions of lowered worth, as though they would surrender to argument, but to learn how to resist implicitly accepting them moment-to-moment. These habits of de-escalation with oneself are what amount to forgiveness when no wrongs have been committed.

But "largely indirect" does not mean "completely." Thus far I have developed a metaethics for the depressed, identifying the form that an ethics for the depressed might take and the problems to which it should be most responsive. The question remains what "ethics as a reliable guide," which accepts that some form of moral luck is ineliminable and so moderates its assignations of credit and blame, may offer at the level of normative ethics. In the context of an ethics for the depressed, to answer this question is also to answer what an ethical theory can do for the depressed even when their central dilemmas involve compulsions unresponsive to argument. Such a theory would be, to borrow a phrase from Philippa Foot, a system of hypothetical imperatives, conditioning its normative claims on already held or endorsed values and so lacking an ultimate justification that, say, any rational agent must accept insofar as they are rational. Yet, even granted this built-in contingency, I am resolved that an ethics for the depressed must have content, not just form. Developing an example of a normative ethics consistent with what I have said so far is likely to be helpful to both professional philosophers evaluating

the merits of the metaethics and to depressed persons seeking a way of ethical thinking that they can tolerate and accept. I will take existential guilt as a starting point for an ethics that those lacking said guilt can accept as properly ethical. In so doing, I will demonstrate to depressed persons that though existential guilt feels like an inescapable trap, it may be turned against itself. Insofar as I succeed at that, I will demonstrate to philosophers that argument, our orthodox method, may be most valuable not for what it tells but for what it shows in its telling.

I begin by situating this effort with an anecdote of ethical action and development that, while humble or petty, shocked me with its success: my own effort to clean up more.

1. Cleaning and Contingent Value Ranking

One day, my wife came home from work and, as I greeted her, I realized how many messes I had left untouched throughout the day while I was writing and, though she did not seem to notice them, I felt ashamed of my lack of conscientiousness. Even if I was working, I could have spared a few moments to take care of them, and even if I didn't make the messes, we share responsibility for our living space. If I want to sustain our mutual trust and avoid defaulting to an unfair status quo in which she likely cleans up more while I remain comfortably unaware, I reasoned, our options are either mutually agreeing on a more fine-grained distribution of chores or my generally being more proactive. After I raised the issue and we discussed it, since she didn't find a different distribution of labor to be necessary, I resolved to try to change my habits.

I knew that doing so would be difficult. First, habits are by their nature intransigent, serving as cognitive shortcuts that require attention and effort to interrupt.

215

Second, I already had framed my habits in a moralized manner, as representing a lack of conscientiousness on my part. Despite the low stakes of this specific problem, if I approached it in a manner likely to fail, I would also be inclined to see that failure as reflecting on my character and so begin to perceive the effort as involving higher stakes, elevating my anxiety and sense of risk out of step with my actual beliefs about the task. I decided to try to come up with a description of my pattern of behavior that was as descriptively accurate and free of normative judgment as possible.

Introspecting, I thought back to a moment in which I could at least vaguely remember passing over some dirt on the floor that I could have swept up in seconds. I tried to recall in what sense I perceived it, that is, what I felt as I saw it. Though I could not be certain that I was remembering rightly, I felt that I had been "propelled" away from the sight of the dirt and that I felt a short burst of relief in the process. Both the content of the experience, the sensation of relief, and the structure, the abruptness of the aversion and the lingering of the relief, seemed to signify that I was seeing the mess *as* a mess: why else look away? This pattern squared with a narrative self-concept I have about myself as the child of an alcoholic. As I feel personally responsible for any crisis that threatens the foundation of my family or way of life, so do I find relief in identifying a potential problem as *not* being a crisis and thus something I can overlook. Primed to see possible responsibilities everywhere, I also long to disavow them and so be released from the tension I constantly generate. I am motivated to look away when I feel I can.

Deciding that this was a likely story behind my behavior, I reaffirmed that I wanted to change. The preceding explanation frames my problematic behavior as actually being a kind of solution, a habit I had developed to manage my own neuroses. But I no

216

longer rely on such experiences of relief. I am no longer powerless in the way that I was as a child and, moreover, I want to affirm that I am no longer so powerless. The satisfaction I will derive from asserting my agency, I concluded, will likely exceed the relief I feel at being swept away by my peculiar survival habits. Yet this satisfaction is derived not just from asserting my agency as such but from doing so in pursuit of what I am endorsing as a higher value: in committing to this change, I am affirming the primacy of conscientiousness over my comfort, at least in the short term. I do reasonably predict that becoming more conscientious will lead to more satisfaction, and hopefully more comfort in some sense, in the long run. But I do not know that for sure, and even if that prediction fails, I would still maintain that conscientiousness is the higher value.

I believed that I was valuing a higher virtue, conscientiousness, over a lower value, my own comfort. Even so, in describing my own pattern of behavior, I avoided the language of virtue or morality. It may be true to say that, if I were to choose my own comfort over becoming more conscientious, I would lack virtue. But it would not be helpful, I think, to bring that up during the process of my becoming more virtuous, not as someone inclined toward self-denigration in a manner that inhibits my capacity to act.

The challenge was, of course, changing myself. I affirmed in principle (or in pride) that I could do so by sheer willpower: every time I felt the aversion to a perceived non-crisis problem like a mess, I could resist and overpower that aversion by recalling my motivation to change myself. But this path seemed needlessly fraught: it would be tiring, and any time my will failed me, I would be likely to spiral into rumination about what such weakness said of me. I decided to sidestep the problem. Instead of forcing myself to clean up messes as I saw them, I would arbitrarily set a date and time, once a

week, to clean up all the messes around. Though this would not be as likely to present my wife with a clean house at any given moment, it would serve as gradual progress toward becoming a more conscientious person with less to be ashamed of.

To my immense surprise, the simple existence of this arbitrary self-imposed obligation quickly transformed my response to messes as I saw them. I would consciously have unbidden thoughts like, "Well, I'm going to clean it up in three days, so why wait?" These non-crisis problems appeared as non-threatening yet unavoidable, so not only did I experience no anxiety about them, I was soon taking the initiative to go above and beyond my self-imposed obligation. Indirectly, I had made great steps toward achieving the kind of self-transformation I had sought. I have largely been able to keep up these habits since and, most importantly, when I slip, I do not feel guilt, only the resolve to return to them. I am confident that I will be able to continue, too, if only because after talking myself and my methods up so much, I had better.

But seriously: the key was that I wanted to clean up. I wanted to become more conscientious, to be a responsible adult, to experience menial tasks as small but heartfelt gestures of kindness toward someone I love. My problem was never a lack of motivation. It was the presence of obstacles, like guilt or self-doubt, undercutting the translation of that motivation into action or long-term planning. That motivation has all the hallmarks of a depressed person. It emerges from a fixation on my own character and the importance of improving it that is probably excessive or even narcissistic. As such, it potentially involves projecting concerns or frustrations onto someone who never voiced any. But I bettered myself and avoided any downward spiral.

Notice how many steps there were in this process. I had an initial intuition, I discussed it with a loved one, I introspected, I analyzed my memories to discern a structure, I compared that structure with a narrative self-concept to identify a likely problem, and I developed an indirect solution to that problem that minimized moralized language or the use of willpower. Each layer is imperfect. My intuition may go astray, my conversation may not be revealing, my introspection may go awry, my memories may be faulty, my self-concept may be inaccurate, and my prospective solution will be based on any or all of those mistakes. I almost certainly got lucky that my solution worked as it did, so it is hard to know to what extent I can generalize from this example. But in a story like this, I see the glimmer of a way forward for a depressed person yearning for ethical growth. I separated one set of intuitions asserting that I could and should become a more conscientious person from another asserting that I was sure to fail in doing so.

In my focus on a change in habits that acknowledges a continuity between traditionally prudential concerns, like feeling more empowered and satisfied, and traditionally moral concerns, like becoming more conscientious, the approach I have described resonates with pragmatic ethics in the style of John Dewey and William James. In my turn toward a form of reflection that identifies the structure of my experience, my approach may involve a kind of phenomenological method. In taking intuitions and careful description of context as my starting point, my approach parallels the pluralism and descriptive ethics of moral psychologists like Owen Flanagan. The normative ethics I argue for in this chapter draws from these three streams. It is an attempt to persuade depressed persons who long for strictness, for a certain and absolute measure of their ethical worth, that a reliable guide can suffice. This attempt will be incomplete until the

219

next chapter, in which I incorporate the lessons of the philosophy of vulnerability, including feminist phenomenology and the ethics of alterity, to discuss the most reliable motive of a depressed person: what Cholbi has called altruism and what I will call "the call of the other." But I begin with a plurality of likely motives, including and especially a desire for reliability.

My normative ethics for the depressed is what I call a value ethics of engagement. While there are many reasonable candidates for what sort of thing a "value" is or can be – a belief, a concept, a type of desire, an object of desire – the definition of value most salient to my value ethics is valuing as a pre-intentional state akin to a second-order quasi-belief or habit. That is, I posit "values as pre-intentional" as being structures of perception, emotional response, initial intuition, and action preparation. In this understanding of value, I am inspired by Max Scheler's phenomenology of value, John Drummond's revised Husserlian value ethics, and Steven Crowell's phenomenology of normativity. I apply the definition of the pre-intentional that I have revised from Matthew Ratcliffe's phenomenological account of depression to show that the same concept, useful descriptively, is also useful in a normative context, that is, in articulating an ethical response to depression. I also note that a similar notion of valuing as pre-intentional may be derived from Dewey's theory of valuation, drawing on his previously discussed theory of habit, thereby offering a naturalized alternative to transcendental (or sometimes "quasi-transcendental") phenomenology. The result of adapting these theories is a framing of ethical intuitions as feelings modified pre-intentionally by one's subjective "value paradigms," the set of which Scheler calls the *ordo amoris* or heart.

220

The terminology of the "value paradigm" comes from Moral Foundations Theory (Graham et al. 2013) and Jonathan Haidt's arguments for social intuitionism, which serve as a prominent example of an empirical account of values understood as shaping intuitions as opposed to as, say, reflectively endorsed beliefs. However, if I am correct in understanding a depressed person as both motivated and undermined by their intuitions, an ethics for the depressed cannot be a straightforward intuitionist normative theory in which a depressed person is simply justified in going with their gut. Even if we can count on depressed persons being consistently altruistic, they will experience those altruistic inclinations as part of a whole or circuit with self-denigrating inclinations that produce contradictions, like clashing beliefs and quasi-beliefs, and inhibit their capacity to act or form intentions. Valerie Tiberius's critiques of Haidt's social intuitionism and Owen Flanagan's comparison of MFT to Mencius's "four sprouts" theory show that both that Haidt's moral dumbfounding arguments for intuitionism fail and that any theory that begins with values as intuition-guiding requires some capacity akin to practical wisdom, like effective agency, to make sense of how value conflict is adjudicated.

Just as an ethics for the depressed, insofar as it is a kind of "ethics as a reliable guide," cannot be a straightforward intuitionism, so can it not follow directly in Scheler's footsteps as an objective nonformal ethics. Despite their differences, phenomenologists of value like Scheler and Drummond, believing their method to uncover necessary and universal structures of consciousness, have only ever sought to ground an impartial ethics on phenomenology. Central to that project is an objective hierarchy of values that may be phenomenologically revealed. Drawing on Drummond's critiques of Scheler and Crowell's critiques of Drummond, I will reject the transcendental method, advocating

instead a critical phenomenology of value that acknowledges the incapacity of phenomenology to offer pure descriptions of structures of consciousness. I see this sort of critical phenomenology as potentially compatible with the naturalism of moral psychology and at least some versions of pragmatic ethics.

But, thus far, these concerns remain primarily formal. The central question of normative content for this approach will be if anything is left of the Schelerian notion of "value ranks" in a partial or reliable ethics developed from a critical phenomenology of value and, thus, what manner of normative guidance I profess this theory to be able to offer. My review of the phenomenology of value, the moral psychology of value paradigms, and pragmatic ethics contextualizes and culminates in a sketch of what I call "contingent value ranks" as the basis for a value ethics of engagement that is also what I call a "gentle perfectionism."

I argue that a value ethics of engagement may involve contingent value ranks, ranked according to their tendency to produce a feeling associated with engagement: a sense of both elation and connectedness that I call a "sense of stability." Elation is a cluster concept referring to the pleasant sense of excitement that can accompany activity that seems worthwhile or significant in the moment, including senses of being present or in the zone. But elation may also feel unstable or excessive, brittle or fragile, and induce uneasiness or confusion or self-doubt even as one is experiencing it. Such a complex of feelings may lead one to, for example, ascribe to oneself a manic state. I use the term "sense of stability" to refer to the cluster of experiences of elation that lack this sense of instability and involve instead a sense of connectedness with a person or purpose, including senses of belonging or being at home. I distinguish engagement from simple

elation by claiming that the experience of engagement, as I use the term, involves a sense of stability. Whether I am engaged with other people or with my own task or project, I have a distinct impression that I am "connected," that the context of my elation has an existence exceeding my own and that I may thus rely upon it or return to it for further experiences of value.[37] Implicit in this experience of self-transcendence and its atmospheric quality is that as the situation is stably valuable, so does it seem to reflect the stability of my own personality: if I sense that I will find these same circumstances valuable again, I am also sensing that I will remain a person who will find them valuable. A depressed person's reliable attraction and sensitivity to a sense of stability is linked to their vulnerability and aversion to senses of precarity and isolation. Stability means stability as a source of experienced value of any value paradigm.

Furthermore, and this is the crucial normative step, I claim that the value paradigms of compassion and fairness are, overall and in the long run, better sources of the sense of stability than the value paradigms of, e.g., loyalty, authority, or sanctity, because the latter imply an in-group/out-group distinction and the former do not. That is, values that are definitionally more partial are in practice less likely to be experienced as stable. If this holds, then insofar as someone may be relied upon to value the sense of stability, they may be relied upon to aim toward the values they predict to be more conducive to that sense, and insofar as they are justified in their predictions, these values are "higher." The result is an ethics of partiality, one that puts the prudential and moral or personal and impersonal on an even footing, that still tends toward the more impartial

---

[37] I am grateful to my therapist, Dr. Brent Horner, for his insight as a psychologist into how best to define a "sense of stability."

value paradigms, achieving what I take to be the closest possible properly justified – that is, acknowledging of moral luck – approximation of an impartial ethics.

I envision this ethics as a gentle and pluralistic perfectionism, that is, as perfectionist in a mostly formal sense. The claim that "fairness is higher than loyalty" is justified in terms of contingent value ranks by appeal to a vision of what would be true of a person's heart or set of values under the best conditions for promoting the sense of stability, that is, a "perfected" heart. This perfectionist commitment is required to retain a difference between what one happens to value and what one should value, broadly speaking. Given the breadth and provisional nature of contingent value ranks, however, a reliable rather than objective value ethics would have a wide sense of what is forgivable or permissible as distinct from what is estimable or virtuous. To act with integrity or good faith, or to act in accord with one's values as pre-intentional while taking value ranks into consideration in deliberation, might be thought of generously as to be on the path of "perfection" or growth. While a virtuous person would oppose, say, cruelty or unfairness toward an out-group from both 1) the person of integrity loyal to an in-group but aware of their inversion of values and 2) the reactionary who rejects higher values and fails to recognize their relevance to deliberation, the assignation of blame to the former would be moderated or gentle. That said, this is a non-ideal theory: though I think some vision of the ideal, like the "perfected" heart, is needed to distinguish what is from what should be, it remains a post hoc extrapolation from concrete experiences of value defined by value paradigms. I thus do not discuss the notion of a gentle perfectionism at length in this chapter and focus instead on values as pre-intentional.

I pause to note my key assumptions and challenges. I am assuming that depressed persons are aware of a set of feelings and intuitions that may be reliably sorted into senses of unstable and stable elation. I am also assuming that depressed persons develop this capacity because of their vulnerability and aversion to senses of precarity and isolation. I am not assuming that depressed persons are correct in their intuitions or assessments of, say, stability, because I am making a point about what depressed persons are inclined to aim for, not what they actually aim for in any given situation. I am also not assuming that I have given a complete account of the sense of stability, just that it involves both elation and a sense of connectedness that is sometimes absent from elation. Most importantly, I am assuming that the value paradigms of compassion and fairness are "overall and in the long run" more conducive to the sense of stability than the value paradigms of loyalty, authority, and sanctity.

This last premise requires the most qualification. I rely here on a conceptual claim and an empirical hypothesis. Conceptually, I claim that the concepts "compassion" and "fairness" map onto intuited value paradigms of compassion and fairness and, definitionally, do not involve an implicit in-group/out-group distinction the way that, say, loyalty does. This implies that an imperative like "Be kind to everyone" or "Be fair to everyone" will seem like a proper use of the concept whereas one like "Be loyal to everyone" will require more explanation or risk seeming like a misuse of the concept. Empirically, I hypothesize that because of this conceptual difference, the values that do not definitionally imply an in-group/out-group distinction will be more likely to produce experiences of stable value.

The most obvious and serious problem with this hypothesis is that of the walled garden: in many cases, arguably even most cases, valuing loyalty and authority and sanctity within a given community will be more likely to produce a sense of stability or belonging than valuing compassion and fairness because forsaking in-group/out-group values can lead to ostracization or shunning and the very precarity and isolation that depressed persons fear. This is the problem of factions (or partisanship, or "tribalism") manifesting in the context of pre-intentional values. Moreover, when viewed in terms of factions, my empirical claim is indistinguishable from what Haidt would identify as an expression of my progressive values: compassion and fairness are higher than loyalty, which is higher than deference to authority and sanctity. Perhaps I am in a political bubble, a progressive walled garden, and my beliefs about what better sustains a sense of stability reflect that more than they do reality. If so, and if an ethics for the depressed fails to justify self-transcendence within a walled garden, so it may fail to justify transcendence of whatever communal or in-group values happen to be ascendent with that walled garden. This would compound the failure of an ethics for the depressed to be a recognizable ethics, that being one that may justify, on a contingent basis, the rejection of one's own prior ranking of values.

I have three responses. First, my approach is a better-justified alternative to an impartial ethics that fails to overcome the challenge of moral luck while claiming that compassion and fairness should be considered higher values. If the impartial ethicist is dissatisfied with an ethics that so clearly risks failing to transcend communal values, the onus is on them to overcome the problem of moral luck. Second, I think this problem should be seen not as a failing of the ethics but rather the point where we reach the limit

of an ethics for the depressed and must begin to consider a "politics for the depressed"

instead, about which I will say more later in the chapter. For an ethics for the depressed,

in which deliberation is usefully discussed largely in terms of rumination and its

alternatives rather than as discourse with other individuals, it may be best to say that

ethics stops and politics begins when the pursuit of self-transcendence becomes the

transcendence of values partial to one's community. Third, I claim that the depressed

person has a sort of ethical secret weapon: the "call of the other" which, combined with

their own inclination toward alienation, assists them in moving beyond walled gardens.

My final chapter will be devoted to describing and evaluating this "call."

The normative ethics of contingent value rankings is a theorization of depressive

self-help practices. I believe that what a depressed person needs in practice is to

somehow cultivate self-trust and self-forgiveness in the face of existential guilt. That

guilt inclines them toward a strict and absolute impartial ethics that could serve as the

ultimate measure of their moral worth. But this is a misdirection of attention. The ethical

tragedy of depression is that persons inclined to spontaneously act altruistically, without

weighing or justifying their actions in terms of ethical obligations, are also compelled in

rumination to obsess over identifying those obligations and justifications. What I want to

demonstrate is the very existence of the attention to their own moral worth, which is

misdirected toward an ultimate measure, may itself be the motivation and justification for

a partial ethics. With trust in their own worth and altruism, a depressed person is liberated

to pursue self-transcending values as sources and structures of a sense of stability. It is

possible to channel the energies of existential guilt, behind demoralization and

hypermoralization, toward higher values by counteracting pre-intentional states that

227

induce dread and self-doubt. It is possible to separate the sets of intuitions that are other-regarding from those that are self-disregarding and so retain spontaneous altruism without compulsive rumination. In this way, a personal ethics, one which starts with one's initial sympathies, may work toward explaining and justifying the importance of impersonal goods, which may be in tension with those initial sympathies. I think that all this is easier to do when a normative theory that competes with impartial ethics is argued for rigorously, so that even though arguments do not alone suffice to ease depression, those seeking reasons find them.

I will end the chapter by considering the prospects of generalizing an ethics for the depressed into a value ethics of engagement. I will conclude that the problem of the walled garden is a much more serious challenge to the theory of contingent value ranks in the absence of depressive tendencies toward altruism and alienation. First, I turn to the phenomenology of value and the notion of value or valuing as a pre-intentional state that I derive from it.[38]

---

[38] I do not engage with possible alternative ethics for the depressed that are neither impartial ethics nor value ethics, but I note here my preliminary concerns with two alternatives: an ethics of conscience and virtue ethics. I resist the language of conscience because, although I claim that a depressed person's other-regarding intuitions are separable from their self-disregarding intuitions, this separation requires the careful cultivation of habits to counteract spiraling, even after which those intuitions may arise as mutually implicated. "Conscience" implies to me a distinctly moral inner voice which, for a person vulnerable to hypermoralization, is a highly fraught and deceptive notion. I also resist the language of virtue because, as noted in my example of cleaning, I think that such language is for depressed persons likely to trigger and deepen hypermoralized deliberation, even though virtue ethics does not require a moral-prudential distinction in decision-making. While virtue-talk may be useful in describing character or identifying some ethical aims (like becoming more conscientious), I think that it is not useful for depressed persons in the practice of building character, which I take to be, ironically, the primary concern of virtue ethics. That is, even in the absence of a moral-prudential distinction, it is easier to develop habits when these habits are not framed as morally significant or associated with concepts of merit or blame, especially when the agent is a compulsive moralizer. As thinking in terms of one's own virtues or character risks hypermoralizing, so does thinking in terms of role models seems to imply a lack of the demoralized alienation common to depressed persons. I believe these points to form a novel normative objection to virtue ethics on its own terms, but I will not elaborate on it further here.

2. Felt Values, or Values as Pre-Intentional

Like the concept of emotion, the concept of value is multivalent and thus difficult to define. The Schwartz Theory of Basic Values (Schwartz 2012), to take one example, defines values as "beliefs linked inextricably to affect" that "refer to desirable goals that motivate action" and "serve as standards or criteria" (3). This is all true, and for the purpose of describing the kinds of values persons hold, it is not necessary to know how a belief can be "linked inextricably" to affect or how it "refers" to motivating desires.[39] But to understand how values motivate, it is clarifying to make a preliminary distinction between values as beliefs and valuing as desiring. A sincere assertion that expresses a value judgment represents beliefs that are identifiable as the values that person holds. In another sense, to say that someone values a goal or purpose is just to say that they desire to achieve it, and so valuing is synonymous with desiring. Yet values "serve as standards or criteria" not just in reflection or conscious deliberation but in shaping one's intuitive or immediate response to a problematic situation. It is because I value loyalty that I am shocked by a perceived betrayal or because I value sanctity that I am revolted by a transgression of taboos. It is this kind of valuing, the kind that guides ethical intuitions, that seems necessarily (or very consistently) affectively charged, not the previous two kinds. Although values are not themselves emotions – "loyalty" and "sanctity" are not like "fear" or "anger" – they seem to ride on emotions in some sense, not just in how they involve affect but in how they structure attention or prepare for action. Some kinds of

---

[39] Schwartz names ten values -- self-direction, stimulation, hedonism, achievement, power, security, conformity, tradition, benevolence, and universalism – grouped under the four headings of openness to change, self-enhancement, conservation, and self-transcendence.

values or valuing may cash out into standard types of beliefs or desires. But the kind of values that guide an individual's initial determination of what is problematic about a problematic situation do not.

I propose that values as intuition-guiding, or "felt values," are pre-intentional. The propositional status of felt values is ambiguous in a similar manner to the conditions of demoralization. A felt value might be understood as a second-order desire with a specific propositional content, a desire to desire in certain ways or to form certain beliefs that are sensitive to the value in question. But a felt value might just as well be understood as a disposition toward sensitivity to that value and thus as a mental state without propositional content. The introspective opacity of felt values, as mental states structuring other mental states, means that this ambiguity will likely not be resolved. Thus, the category of the pre-intentional, defined in terms of introspective opacity and functional ambiguity regarding propositional status, best accounts for their intuition-guiding and affect-shaping characteristics.

The previous discussion of demoralization shows how a pre-intentional state may lead to the loss of motivation: a second-order quasi-belief or habit of feeling may contradict a belief or suppress a desire and disrupt the passage of motivation into action even while motivational states remain. I think of felt values as the positive correlates of such states, promoting rather than undermining action. The analysis of depression, in developing the definition and demonstrating the usefulness of the category of the pre-intentional, clarifies how best to understand felt values and distinguish them from values as beliefs or first-order desires. Moreover, depressed persons concerned with the conditions of their own motivation or action may be especially sensitive to the operation

of felt values on feelings associated with motivation like elation or demotivation like anxiety. These feelings present themselves as value-laden, as representing a lifeline-like lingering attachment to an alienating world or the threat of losing this attachment. Yet although values may represent ways of caring or reasons for living, value-talk is depersonalized: to say that one is "sensitive to loyalty" is not equivalent to the character assessment that one is "loyal." The expectation that depressed persons will be attuned to the felt values guiding their initial intuitions and will be able to think about them without spiraling into negative self-assessment justifies turning to a value ethics for an ethics for the depressed.

As with the account of depression, the account of felt values relies on phenomenology: felt values are a structural element of consciousness. Felt values are not felt as feelings, as though a "feeling of loyalty" were to accompany a feeling of anger at betrayal, but as implicit in the seeming "direction" or temporal and attentive structure of feelings. It is thus unsurprising that the closest analogue to the account of felt values I have presented is found in Max Scheler's early twentieth-century phenomenology of value and his contemporary commentators and Husserlian critics. I also use an introspective approach analogous to phenomenology, which I think of as a form of critical phenomenology, in my cleaning example when I reflect on the structure of a past experience. But where my approach is fallible and tentative, the Husserlian tradition of phenomenology relies on a transcendental method that posits necessary structures of consciousness and Scheler's (1973b) variant is defined by a phenomenological "attitude of spiritual seeing" that ostensibly reveals an objective ranking of values. The transcendental method of phenomenology, specifically the transcendental reduction

following the *epoché*, seems to me still insufficiently justified given that it aims to make claims of necessity. But my primary concern with the method here pertains to developing an ethics for the depressed. Depressed persons are ambivalent toward their intuitions: some intuitions present themselves as sources of much-needed motivation and others as threats that suppress motivation. Such persons depend on some practice that helps to sort intuitions as trustworthy or untrustworthy. But this practice will always be imperfect and its conclusions therefore hedging. Overconfidence in a method of analyzing or deconstructing intuitions will strike a depressed person as not just mistaken but dangerous. This is how transcendental phenomenology has always struck me.

Within the phenomenological tradition, as Drummond's (2021) detailed account shows, phenomenology of value has accompanied attempts to establish moral objectivity and an impartial ethics, including Drummond's own. To do a phenomenology of value for an ethics for the depressed will be to distinguish the useful content of earlier phenomenology from its approach, which finds in the apprehension of transcendental structures of experience some necessary and thus universal grounding for a normative ethics. Drummond convincingly critiques what he calls the "strong value realism" of Scheler and defends a modified Husserlian account of values. But Drummond makes two central errors: 1) even after modifying Husserl's theory of presentation, he presumes that there is a purely descriptive moment in perception in which we see a thing just as it is, a moment treated as phenomenologically distinct even if supposedly not "separable" from evaluative perception, and 2) he conflates the causal and the justificatory in his discussion of this moment, relying on an underexamined notion of "prima facie, non-inferential, and defeasible justification" or felt warrant in establishing the objectivity of values. The first

232

error is what Shaun Gallagher calls a confusion of conceptual and phenomenological distinctions and resembles what Elizabeth Anscombe might call a "brute facts" problem of how to determine the bounds of a description besides normatively. The second is what Steven Crowell (2002), in a Kantian vein, calls a problem of heteronomy in which agents, being guaranteed to have certain values in what is supposed to be a personalistic ethics, do not have the opportunity to endorse or reject those values, thus making reasons into causes.

Crowell (2013) argues that phenomenology is a fundamentally normative discipline, despite its positioning as descriptive, because it is about "not things but the meaning or intelligibility of things" (10), which involves conditions of correctness or appropriateness. His interpretations of figures like Emmanuel Levinas and Martin Heidegger are illuminating in showing that what might look like ethical ideas with insufficient content, like the "face" or the "call to conscience," are attempts to account for the emergence of this experienced normativity and the subjective grounds of responsiveness or responsibility. But Crowell remains seemingly agnostic about the possibility of the transcendental reduction and unclear if phenomena like the "face" or "call," which operate akin to the transcendental reduction, are episodes that a subject may encounter or have (in the Heideggerian's favorite phrase) always already been encountered in some sense. I venture that this is because the notion of a pure phenomenological description is in tension with his account of intentionality as normative because phenomenology requires a phenomenologist to attend to a phenomenon, which, if Crowell is correct, would be a process structured by some experienced normativity. His phenomenology seems methodologically trapped in a

233

performative contradiction, and in his account of how one becomes an object to oneself, he neglects the phenomenology of vulnerability, to which I will turn in the next chapter.

I proceed to summarize Scheler's (1973a) value theory and the concept of the heart or *ordo amoris*, to which I will take MFT's value paradigms to be a naturalized analogue, and of Schelerian love as the movement between value ranks, which I will go on to compare to Dewey's concept of growth. Through my review of Drummond's and Crowell's accounts of phenomenology and normativity, I will sketch out a pragmatic phenomenology, one that may perform a naturalized version of the *epoché* but not the transcendental reduction, as the basis of a phenomenological pragmatism informed by the phenomenological tradition in its development of an introspective practice to accompany deliberation.[40] A reliable value ethics calls for a phenomenology of value in the style of contemporary critical phenomenology, taking what lessons it can from classical transcendental phenomenology.[41]

The two central claims of Scheler's value theory are that values are independent of their bearers and that feelings are intentionally directed toward values. A "feeling-

---

[40] In this, I am inspired by Natalie Depraz et al.'s (2003) work in neurophenomenology and Claire Petitmengin's (2006) practice of microphenomenology, which adapt the *epoché* to an empirical psychological context.

[41] For reasons of space and relevance, I refrain from discussing William Hosner Smith's 2012 *Phenomenology of Moral Normativity*. Smith argues that a full account of the bindingness of morality requires aspects of two approaches: the first-personal, for which the bindingness of morality depends upon the agent's practical deliberation or reflective endorsement of norms (as in Korsgaard's Kantianism or Drummond's Husserlianism), and the second-personal, which grounds morality in the demandingness of face-to-face encounters with other persons (as in Darwall's account of dignity or Levinas's account of the face). In Smith's two-pronged approach, the foundation of morality as such is second-personal and the bindingness of moral obligations is established first-personally through situated and deliberative endorsement. His project stands in stark contrast to my own: he commits fully to transcendental phenomenology, he commits to the objectivity and overridingness of moral obligations over prudential interests, he is concerned with the justification of norms to the exclusion of the apprehension of values, and he commits to grounding morality in face-to-face encounters where, in Chapter 6, I argue that this is a limitation of ethical possibility that reinstantiates the restrictions of existential guilt. While Scheler and Drummond also argue for impartial ethics, their phenomenology is at least a phenomenology of value like mine.

state" like pain does not in itself have intentional directness. For Scheler, however, enduring pain or suffering pain are "feelings," ways of experiencing a feeling-state differentiated by the value, like fortitude, that is its intentional object: "Pain observed is almost the opposite of pain suffered" (1973a, 256). Because a feeling directed toward fortitude is not itself fortitude, just as a durable object is not itself durability, so Scheler claims that values are independent of their bearers. Moreover, the apprehension of a value precedes that of its bearer: the value "is the first 'messenger'" of its bearer's "particular nature. An object may be vague and unclear while its value is already distinct and clear" (18). These claims are central to Scheler's nonformal value ethics and his criticism of Kant. For Kant, ethics must be formal, based on a universal prescription like the categorical imperative that is not responsive to an individual's values because, he argues, nonformal ethics is contingent upon sensory perception of goods in the world and so vulnerable to skepticism. Scheler affirms that ethics requires an objective *a priori* foundation but counters that values are apprehended prior to goods in Kant's sense and thus that it is possible to develop an objective ethics that centers the person and their values rather than disregarding them as contingent. Scheler's nonformal ethics depends upon his value theory showing that an objective ranking of values may be apprehended with certainty even though this ranking is not formally necessitated.

Scheler thus distinguishes between subjective value ranking and objective value ranks. All values have a "height" that is given in acts of preferring or "placing-after," which Scheler claims is distinct from, prior to, and necessary for willing or representational thinking about a value. To prefer a value is to gravitate toward it or be oriented toward it in a way that makes it available as, say, an option to be chosen. For

Scheler, it is because one already prefers fortitude in this sense over, say, fragility that

one may exert one's will to endure pain rather than suffer it. While "the '*ordered ranks of*

values' are themselves absolutely *invariable*," the "'rules of preferring' are, in principle,

variable throughout history" (88). Thus, individuals may undergo a subjective "*widening*

of the value-range" that opens them to the apprehension of objectively higher values or a

"*deception of preferring*" that precedes the choosing of "an end founded in a lower

value" (88-89). Scheler distinguishes between four value ranks: from lowest to highest,

the sensible, associated with the pleasant or the agreeable; the vital, understood in terms

of the noble and the vulgar and linked to strength or ineptitude; the spiritual or mental,

including the aesthetic, the juridical, and the epistemic; and the holy, given through

worship and repentance and evocative of bliss or, through absence, despair. Good and

evil are not themselves values but occur in conjunction with, or "ride on," acts that

realize higher or lower values. Scheler offers what he calls essential or rational criteria

for ranking values, including divisibility, duration, degree of fulfillment, and

absoluteness, all of which place unity above division and eternity above transience. But,

crucially, though Scheler calls these criteria essential, they function in this philosophy

merely as guidelines for value ranks that must be grasped through feeling for Scheler's

ethics to count as nonformal.

The subjective widening or narrowing of values and value ranks that one

apprehends is possible for Scheler because of love and hate. Love is "a movement in

whose execution ever *new* and *higher* values flash out, i.e., values that were wholly

unknown to the being concerned" (261). Love makes new and objectively higher values

into possible intentional objects of feeling for a subject. But Scheler also understands

236

hatred as "a *positive* preoccupation with lower possibilities of value" rather than a negative turning away from higher values (1954, 153). This, and his distinction between preferring and intentional feeling made possible by preferring, grounds his interesting account of *ressentiment*. For Scheler, *ressentiment* cannot be defined by repressed feelings like envy or spite, because to need to repress these feelings implies that, at the level of preferring, one is still valuing the values at which such feelings are intentionally directed. True *ressentiment* inverts preferring, leading higher values to be fully apprehended by the subject as lower, such that those others who are directed toward higher values "cease to be enviable, hateful, and worthy of revenge. They are unfortunate and to be pitied, for they are best with 'evils'" (1998, 56-57). The subject of *ressentiment* is entrapped in a deep "organic mendacity": they no longer need to lie to themselves or repress anything because their pettiness and vulgarity are sincere, which is to say that their behavior is consistent with their real preferences, a "systematic perversion and reinterpretation of the values *themselves*" that precedes one's beliefs or conscious value judgments (57). *Ressentiment* is a twisting of the heart or *ordo amoris*, one's particular rules of preferring or felt priority of values. Scheler claims that "*our heart is primarily destined to love*, not to hate" and that "every act of hate is founded on an act of love," a confused defense of the values dear to one's heart, "without which it would lack all sense" (1973b, 125-126). But *ressentiment* remains a possibility that may close off higher values to a subject and prevent them from apprehending the objective order of values through love.

This is a serious problem for Scheler because the possibility of a nonformal value ethics depends on a subject necessarily being able to apprehend the objective order of

values, which hate and *ressentiment* would prevent. If phenomenology is how this objective order is identified, then phenomenology must be able to transcend subjective narrowness. Scheler thus rejects Husserl's *epoché* as a methodological foundation for phenomenology. The *epoché* is an act of judgment that suspends the "natural attitude" of everyday life, in which subjects experience the world as really existing as it is experienced, unaware of how the world as it is experienced is shaped or constituted by the structures of their consciousness. The *epoché* is thus the preliminary move in turning toward the constituting structures of consciousness through the transcendental reduction. But for Scheler, an act of judgment is an inadequate ground for a phenomenological investigation into value because preferring precedes and structures judgment. Scheler's more radical phenomenological "attitude of spiritual seeing" is not a method or "goal-directed procedure" (1973b, 137). But Scheler only gestures at what it entails, and though he is aware of "phenomenological controversy" and claims that it "is not beyond settlement," he does not say how (155). Manfred Frings (1997), in his commentary on Scheler, states that these "psychic techniques" are "similar to those in Buddhism" and involve a "purification of spirit" (191).

Scheler's value theory inspires my understanding of felt values. His counterintuitive claim that feelings like pain may have an intentional object detectable from the way in which they are experienced grounds his theory of pre-judgmental preferring and of the heart as the set of a subject's preferences, all of which resonates with the category of the pre-intentional that is central to an ethics for the depressed. Felt values are akin to Scheler's values as the objects of feeling, and the valuing of felt values is what Scheler calls preferring. But Scheler's view of phenomenology as a practice is

underdetermined, making him vulnerable to a key objection: a phenomenologist ostensibly identifying the objective order of values might be afflicted by *ressentiment* as he defines it and so unwittingly present an inverted hierarchy. The Schelerian phenomenological attitude may be analogous to meditation, but meditation is a practice that can succeed or fail. Moreover, if one knows that one succeeds or fails at meditation by a feeling, say of tranquility or disquiet, then one's practice is dependent on a feeling that, like all feelings, may be perverted by *ressentiment*. Perhaps Scheler himself was afflicted by *ressentiment* and so wrongly apprehended the priority of the eternal over the transient in his phenomenology of value. There is no recognition or discussion of this problem in the available commentary on Scheler, but I speculate that Scheler's confidence in the openness of a phenomenologist's heart to an objective order of values follows from his faith in God and the possibility of repentance, to which we are roused by "God's love, constantly knocking at the door of the soul" (1972, 65).[42] Regardless, absent a capacity for "spiritual seeing," it is impossible to justify the claim that a subject has access to an objective order of values and thus that Schelerian value ethics is an equally impartial alternative to Kantian ethics.

Drummond's (2021) review of Scheler's "strong value realism" develops a critique from Dietrich von Hildebrand and Nicolai Hartmann: there is a tension between

---

[42] I discuss this account of Scheler's phenomenological attitude at greater length in "Scheler's Phenomenology of Value as Value Pluralism" (Fitzpatrick 2022). As Crowell (2002) notes, ethical personalism is often rooted in religious conceptions of the person. In naturalizing my understanding of value paradigms in the next section, I am implicitly situating my phenomenology of value as agnostic toward God. I am also hesitant to identify the heart as the set of value paradigms with the person as, say, the bearer of political rights, which I take religious personalism to be willing and preparing to endorse. In general, I am attempting to avoid any naturalized teleology. To be able to see one's values and purposes reflected back at oneself in the world, and to start ethical deliberation there, is not to actually see one's values and purposes in the world as found or already justified prior to deliberation.

the universality of ideal or objective values as proto-normative and the universality, or categorical status, of obligation as normative. If the apprehension of an objective value is the basis for the willing of an imperative or obligation, the latter "appears as an imperative only to those who recognize the value as desirable or lovable" (128), which is to say that it is hypothetical, not categorical. If "moral imperatives grounded in ideal values should have" a "categorical character," then ideal values must somehow be "experienced as universal demands." Yet the flip side of this tension is that a universal demand, in not being relative to the contingencies of one's heart or values, implies that "an agent is replaceable by anyone, and this is to deny the agent's individuality as a person." Personalism and impartiality seem intrinsically at odds. Drummond further complicates the relation between values and obligations by noting that, contra strong value realism, valuing is not sufficient for motivation: "The relation between valuation and action is mediated by desire" (129). Though I think it more accurate to say that the relation between desire and action is mediated by valuing as pre-intentional, Drummond is right to note that desire is missing from this account of moral psychology and that value ethics is in tension with categorical imperatives. Not just the objectivity but the normativity of phenomenological value ethics, its status as rule-issuing and action-governing, is threatened by the challenges of reconciling universal values with universal obligations and of its potential subordination to sentiment in guiding action.

The central normative question is what it takes for a value to count as a norm. The virtue of Kantian ethics is in how its concept of self-legislation squares objectivity and autonomy: reasons for action are necessitated by rationality itself, not contingent on sentiment, and are therefore universally binding, but also rationally endorsed by each

individual, thereby preserving a categorical distinction between reasons for action (or freedom and the normative) and causes of action (or nature-as-mechanistic and the descriptive). The weakness of value ethics as an impartial ethics seems to be heteronomy. If valuing simply moves one to action, the individual does not choose and values as "reasons" are simply causes, throwing into question the basis of ethics itself. But if choosing to adopt rules based on one's personal values implies the right to reject rules that do not represent one's personal values, then such an ethics is not impartial.

Drummond's reading of the history of phenomenology of value and its personalistic ethics, framed as a transcendental phenomenological alternative to Kant's transcendental philosophical ethics, parallels the previous discussion of psychological realism. The concern that depersonalizing ethics to render it impartial will undercut the sources of motivation for normatively governed action, whether Williams's "ground project" or Wong's "relationships," shows up for Drummond as the question of how to incorporate into ethics the "will to flourish":

> obedience to the moral imperative apart from any reference to inclinations depersonalizes the action—whether in Kant, von Hildebrand, or Levinas—insofar as the action is divorced even from the agent's will to flourish precisely as a moral agent through obedience to the moral imperative and in fulfilling her own moral commitments. The will to flourish is entirely displaced in Kant by obedience to law, in von Hildebrand by obedience to the call of the important-in-itself (value), and in Levinas by the presence of the Other. The question arises whether this is satisfactory as an account of moral motivation and, by extension, of moral normativity for individual agents. (2021, 132)

Drummond attempts to develop a phenomenological normative ethics by establishing a transcendental good of truthfulness. He starts with the Husserlian concept of a "value-perception" that involves both "the cognition of a set of underlying non-axiological properties," its features irrespective of how it is evaluated that it made it recognizable as

the same object under different evaluations, that "serve of motivate the feeling (or emotion) apprehending the object as valued, as good or bad" (133). Axiological and non-axiological properties appear in the natural attitude as properties of the object. Distinguishing them phenomenologically reveals how valuations are "affective responses to non-axiological properties" (2009, 369). Drummond claims that an essential feature of intentionality is "the teleological directedness of an empty intention to a full intention": an agent strives to make intentional objects more fully present to consciousness by, say, seeking out an experience that has only been imagined. He takes this to ground the claim that "reason" is "the striving for evidence" in the Husserlian sense, as the experienced "agreement" of emptier intentions like imaginings with fuller intentions (2021, 134). This is the "transcendental good of self-responsibility," which begins in truth-seeking and ends when "the flourishing agent … necessarily realizes the goods of thinking well, feeling well, and acting well," that is, "the goods *of* rational agency" (135). Because intentionality intrinsically involves the search for evidence in the phenomenological sense, agents are guided by the pursuit of truth to, for example, question if their valuations are "warranted" by their objects (2009, 373).

But Drummond's theory of phenomenological ethics remains equally vulnerable to the challenge of heteronomy, as pointed out by Crowell two decades prior. There are at least three other serious problems that I can only gloss here. First, Drummond's claim that the axiological/non-axiological distinction is given in experience rather than, say, inferred reflectively faces an objection in the vein of Anscombe's (1981) critique of brute facts. Drummond (2013) revises Husserl's original view of the non-axiological (or presentational) and axiological (or affective) dimensions of intentional experience to

emphasize that "the affective dimension does not belong exclusively to the subject" in the sense that the world is presented as "affectively charged" (245). But it is only possible to ground normativity broadly in evidence-seeking if subjective valuations are responses to objective non-axiological properties, as Drummond claims. The set of non-axiological properties of an object is just a description of that object, and descriptions are bounded by concerns of relevance, which is in turn relative to a normative notion of usefulness or purpose. The notion of a distinct dimension of perception that perceives objects "in themselves" is implausible and unnecessary for inferring that an object is the same object under different evaluations. The latter point is an example of a second problem, that Drummond is, as Gallagher cautions against, treating conceptual distinctions that are the product of reflection as being phenomenological or given. There is no way to clearly distinguish the "non-axiological" features of an object through phenomenological reflection because an object "in itself" is simply not given in experience. Third, Drummond's notion of the truthfulness of emotions relies on a concept of felt "warrant" which he identifies as "prima facie, non-inferential, and defeasible justification" (2009, 375). This underexamined notion risks conflating experiencing some feature of the world as like a reason to act with taking it as a reason to act, or in other words, conflating motivation and justification. All of this leads to grounding normativity in a "transcendental good" that, though framed as the good of agency itself, is not itself chosen: it is a cause, not a reason. The answer to the question "why seek evidence?" is, in effect, that we cannot do otherwise. This is heteronomy, just as Drummond points out in the case of Scheler. It is not an account of reason-giving because it is not an account of reasons distinct from causes.

243

Crowell's neo-Kantian critique of Drummond's quasi-Aristotelian attempt to ground normativity in a transcendentalized teleology results in a competing approach to phenomenology and normativity: self-responsibility, not evidence-seeking, is the foundation of phenomenology because it is an account of intentionality or meaning, not consciousness as such. Drummond (2002) claims that the "phenomenological axiologist" may ground obligations in "the experience of the *transcendental goods* of authenticity" or evidence-seeking as "satisfying our *transcendental* interest in the authentic life and as desirable in themselves and for everyone," thereby surpassing the merely empirical and so contingently established teleology of Aristotle (44). He later positions his view as a "weak value realism" (2021, 133). But Crowell points out that, to the Kantian, any value realism "represents a metaphysical (dogmatic) solution to the problem of obligation: by declaring some things to be 'intrinsically normative,' it stanches the potential regress in demanding reasons for why I ought to do something." This disagreement over normativity "mirrors," Crowell says, the epistemological dispute between Husserl and neo-Kantians over if a purely descriptive practice can ground knowledge claims "without any principle for deciding whether these claims are ultimately valid" (2002, 54).

Crowell is concerned to, like Kant, "find the ground of obligation in freedom itself," not in "the normative force of some objective value" (55). But he seems sympathetic to the Husserlian and Schelerian critique that "the value of rationality is itself but one (albeit very important) objective value" that is experienced in "affective intentionality." Their personalism leads to a perfectionist account in which the source of normativity is not the individual heart as such, which is flawed, but the heart as it would be under perfect conditions, guided by love. This can only result, for Crowell, in "a

244

historical vision anchored in something like a personalistic theology or rational faith" (57), which remains a metaphysical account. The alternative is Levinasian: respect for persons is grounded not in "the symmetry of a shared rationality" or a metaphysical account of how one would treat persons under perfect loving conditions, but in "the asymmetry of the other's (ethical) superiority to my (ontological) freedom" (66). This approach takes a primordial experience of responsiveness to be how reason-giving is possible.

Crowell (2013) reads Heidegger, in the concept of the "call to conscience," and the later Husserl as offering such an account of the source of normativity in responsiveness as such. In his account, phenomenology and neo-Kantianism split over Kant's failure to explain how "transcendental knowledge" is possible, that is, how the transcendental schema may be known except through the transcendental schema: phenomenologists turn to a broader investigation of experience while neo-Kantians turn away from experience toward transcendental argument. Though Crowell still sees phenomenology as offering "'presuppositionless' description" (18) of experience, the account of transcendental consciousness "thematized under the reduction" (25) shows that "what I encounter in the world can be held up to norms or standards only because in my very being I must hold myself to standards, that is, understand myself as being something that can succeed or fail" (28). The ground of phenomenology becomes a norm of "ultimate philosophical self-responsibility" (76): because the transcendental reduction cannot involve bracketing "my commitment to being a philosopher" without "thereby losing the very topic of inquiry," phenomenology "must become existential" (77). In this context, Heidegger's account of care and anxiety in *Being and Time* is not just an account

245

of intentionality as disclosed phenomenologically but of normativity. Through the call to conscience, "a creature who is 'grounded' by social norms" is transformed into "one who 'grounds' norms by giving grounds, that is, reasons" (187). In an anxious breakdown, "what looks like a collapse of everything that matters instead reveals the condition for the possibility that anything can matter at all," which is not just rationality, the capacity for reflection, or consciousness as such but the structure of care, including affectivity and know-how, and how meaning arises for intentional consciousness (190). What Crowell calls the "liminal condition" of anxiety and the call to conscience is "the possibility of everyday discourse" or reason-giving (213). Investigating this liminal condition is thus phenomenology's contribution to a complete account of normativity.

Crowell's account of phenomenology and normativity carefully observes the distinction between reasons and causes or justification and motivation, but he does not explain how to understand the method of transcendental phenomenology as an uncovering of the transcendental sources of normativity, and I do not think he can. He seems to acknowledge that the reduction is incomplete in that it does not bracket the phenomenologist doing phenomenology, but uncovering responsiveness as such requires a complete reduction, for which the call to conscience is an analogue. This is a problem for phenomenology that he does not attempt to solve. In later commentary, he states that the anxious breakdown "serves as a *methodological norm* for one who is trying to be a philosopher phenomenologically" (2016, 249). This is consistent with his Heideggerian account, but it is also absurd. A breakdown is an anti-method: it is the absence of rules or guidelines. The call to conscience cannot involve any rule-guided content if it is to be an encounter with responsiveness as such, in the most general possible sense. This reveals

246

an ambiguity in phenomena like the "face" or the "call," which Crowell (2012) takes to be analogous, that is sharpened by the framing of the anxious breakdown. A breakdown is an episode. Yet if responsiveness as such, "the possibility of everyday discourse," precedes the episode, then the call itself is just how one becomes aware of that responsiveness, so it cannot be true that discourse is made possible by the call to conscience. Even if Crowell concedes this, the problem is that doing phenomenology has become something done to the phenomenologist: the structure of care is laid bare when one has a breakdown. This means, for example, that phenomenological conclusions can only be verified, if such a term is even appropriate, by someone having a breakdown for themselves. The very logic that establishes the project of phenomenology, as necessary for accounting for the transcendental subjective conditions of normativity, undermines the viability of phenomenology as the means of pursuing its project.

Fundamentally, if Crowell is correct that intentionality is normative, then when a phenomenologist is directed to the structures of consciousness, this investigation is also normatively structured. This is the same problem as the inability of the phenomenologist to bracket their own doing of phenomenology, just framed differently, but it clarifies the key issue: phenomenology, in uncovering responsiveness as such, uncovers that a presuppositionless act of pure description is impossible, yet this is precisely what phenomenology is supposed to be. For Crowell, the transcendental reduction is a performative contradiction.

Crowell is sensitive to the problem of how to redefine the transcendental reduction when the phenomenologist cannot bracket their own doing of phenomenology, as shown by his "uneasiness" at, and thoughtful account of, Eugen Fink's embrace of

paradox in ostensible defense of the transcendental reduction (2001, 257). But thus far, he

has not raised it as a problem directly or attempted to resolve it.[43] As I see it, the essential

problem for the transcendental reduction, as the turn toward the necessary structures of

consciousness after the suspension of the natural attitude, lies in how the

phenomenologist knows which structures of consciousness are the necessary ones. If

phenomenology gives up its transcendental method, it may become merely a kind of

imperfect variation on introspection. But other methods conducive to phenomenological

accounts of normativity are available, such as Bonnie Mann's (2018) Beauvoirian critical

phenomenology of "oscillation" between concreteness and abstraction. It is possible to

reconceive phenomenology as imperfect or impure and the phenomenologist as fallible

and vulnerable and still retain its object as the structures of intentional consciousness.[44]

Crowell believes that an account of normativity is incomplete without some

understanding of the individual's existential responsiveness to norms. Whether or not he

---

[43] One of Crowell's (2013) goals is to establish that the later Husserl, like the early Heidegger, views phenomenology as no longer a theoretical and apodictically certain science of consciousness but as a practical or normative account of intentionality as care. I find Crowell's account persuasive, but I would counter that there is less daylight between Fink and Heidegger than Crowell might think. As the early Heidegger replaced the transcendental reduction with the call to conscience, turning something one does into something one undergoes, so the later Heidegger, dismayed by modernity (and, I think, the failure of the Nazi reaction against modernity that he endorsed), proclaimed that "only a god can save us now," awaiting rescue by some force beyond everyday intelligibility (1981). Unless phenomenology is reconceived as a critical practice and the phenomenologist is reconceived as vulnerable, the attempt to uncover the structures of experience and the sources of meaning will fall into waiting for someone or something to do it for us.

[44] Edith Stein (1989) presents an alternative early phenomenological approach to felt values. She argues that the phenomenological reduction cannot suspend sensations and that sensations have locations relative to the "zero point of orientation" of the pure "I." The passage from awareness of one's body as lived body to the awareness of one's body as physical body involves the transcendental empathic realization that the "zero point" of the pure "I" is one point among many. Thus, empathy plays a fundamental role in the awareness of one's own body as an object in the world and thereby an object for others, and in the production of the psycho-physical person that is the object of psychological inquiry. Stein further distinguishes the psycho-physical person from the spiritual person possessed of a hierarchy of felt values, who through empathy may discover ranges of value otherwise closed off. Though she initially understands values and feelings as following a rational lawfulness akin to physical causality, she later (2000) modifies her view into a more personalistic approach, though does not attempt to catalogue the values as Scheler does. I rely on Scheler due to the useful parallel between his value ranks and Haidt's value paradigms.

is right depends on what philosophers want from an account of normativity. His position

recalls the debate between Rorty and Derrida on the need for a transcendental or "quasi-

transcendental" account of the conditions of the possibility of responsibility like "infinite

responsibility." Where Rorty argues such a project can only either be trivial or

therapeutic and so non-argumentative and non-philosophical, Derrida emphasizes that the

moment of "decision" at which normative discourse is closed, even provisionally, is a

kind of leap (Mouffe 1996). Rorty's position more carefully maintains the distinction

between reasons and causes. But I think that an ethics for the depressed, in showing the

exhortative effects and context of argument, may show the way to a reconciliation of

Rorty and Derrida's views of philosophy, which I take Crowell, who takes the reason-

cause distinction seriously, to be aiming for as well. Here I will just note that if Crowell

is right that an account of normativity should include its subjective transcendental

conditions, this will be so even if phenomenology can only offer incomplete or fallible

accounts of them.[45]

 This overview of the phenomenology of value and normativity shows an ethics

for the depressed what it takes for a value to count as a norm: the exercise of effective

agency in using a value as a rule for action. A felt value is not intrinsically normative or

---

[45] Derrida (1992) claims that "deconstruction is justice" (15), approaching justice in a quasi-transcendental register: law is deconstructable and must be deconstructed in pursuit of justice, but justice is the un-deconstructable condition of the possibility of deconstruction. Justice is paradoxically an experience of what we are unable to experience (mystical, impossible, an aporia) or of the undecidable, explicitly drawn from Levinas in *Totality and Infinity* as arising from an encounter with the singularity of the other. Critchley (1999) points out that Levinas transitions from an ethical to a political understanding of justice: where in *Totality and Infinity* justice functions as a kind of synonym for the ethical, in *Otherwise than Being* the question of justice arises when a third party enters and obliges one to choose between competing ethical claims. In Critchley's summations: "to be just is to recognize one's infinite responsibility before the singular other as something over which one cannot ultimately decide," and this "propels one forward into politics … from undecidability to the decision" (36). Rorty's response is stark: "I do not see the point of defining a commonly used term such as 'justice' as the name of an impossibility … It seems to me pointless hype" (1996, 43-44).

objectively good. Like a first-order desire that it structures, a felt value begins as a cause, not a reason, and is not necessarily aligned with higher values. Acknowledging this and abandoning absolute impartiality maintains the distinction in a value ethics between motivation and justification and avoids a heteronomous reduction of normativity to pre-intentional valuing. My hope is that the emphasis on affective intuition in deliberation keeps a depressed person in touch with, and broadly legitimizes, plural motivational channels that aim toward a sense of stability. These motivational channels are not just defined by different desires or goals but by the different values that structure them. However, without transcendental access to an objective account of values, it is not clear what these different values are, let alone which are seen as "higher." A reliable ethics thus turns to an empirical account of values as intuition-guiding: the concept of value paradigms.

3. Value Paradigms and Naturalized Pluralism

The use of a felt value as a rule for action or norm begins as subpersonal and is only subsequently conceptualized for use in normative discourse. Consider the example of a third party who witnesses someone betray a friend for the sake of fairness to a stranger by, say, awarding the stranger some gift or position desired by the friend, due to need or merit. The third party's intuitive reaction will be mediated by their subjective ranking of fairness and loyalty: they may feel impressed, indignant, or conflicted about what they have witnessed. As pre-intentional values, fairness and loyalty model the structure of what this person picks out about the situation, how they feel about it, and how they are first predisposed to respond. But fairness and loyalty do not remain just pre-

intentional. These values are named or conceptualized in and for normative deliberation, whether to privately make sense of any internal conflict or to interpersonally discuss their experience with someone else who may disagree.

Of course, the individual most likely relies on received norms that seem to suit their values in deliberation rather than inventing them freshly by conceptualizing their values. The value ethical claim is that even these norms get their "grip" on the individual's intuitions or "leap to mind" in a way structured by values as pre-intentional. Because of this, deep disagreement may arise when one party only "sees" in the above situation fairness or, alternatively, betrayal but not fairness. In turn, a full resolution may require a change in pre-intentional states, involving more than a change in beliefs or desires and likely including changes in habits or dispositions. But, because the pre-intentional structures intuitions, the person so changing might not find the imperative for this change to itself be intuitive. Understanding values as pre-intentional, then, strengthens the importance of normative argument as the means of advocating why someone should willingly attempt to incorporate a value into deliberation that they are not intuitively inclined to notice at stake. To see some norms as downstream of values, insofar as the former conceptualize or grip the individual through the latter, is not therefore to diminish their significance. It may, however, recontextualize argument as one dimension of a broader exhortative project that is important on the basis of a value: even when it fails, in its appeal to the rationality of the interlocutor, it is fairer than the alternatives.

This account of moral disagreement as arising from conflicting value intuitions resembles Jonathan Haidt's (2012) account of political divisions between liberals and

conservatives in *The Righteous Mind*, and the account of value upon which it relies resembles the concept of value paradigms derived from the research of Haidt and his collaborators into Moral Foundations Theory (Graham et al 2013). But Moral Foundations Theory, or MFT, is strongly intuitionist, claiming that "moral evaluations generally occur rapidly and automatically" and that "moral reasoning is done primarily for socially strategic purposes" (66). An ethics for the depressed depends on a general capacity for self-transformative and not merely strategic deliberation because, as I have argued, of the need to discriminate between other-regarding and self-disregarding intuitions. There are also empirical challenges to MFT: though MFT specifies five value paradigms, Joshua Greene (2013) argues the data shows two clusters of care-fairness and loyalty-authority-sanctity (386), and a later study has challenged the global applicability of MFT's specified paradigms (Iurino and Saucier 2020). Though MFT aims to incorporate Richard Shweder's anthropological claim for "three moral languages" of autonomy, community, and divinity, MFT may be unnecessarily subdividing the ethics of community and divinity, thereby weighting the number of values toward those associated with conservatism. But even if MFT's account of value paradigms is partial, it serves as the most prominent example of a naturalized take on values specifically as pre-intentional or intuition-guiding. As Flanagan's critique of MFT's intuitionism shows, a value ethics, like a virtue ethics, relies on some capacity of practical wisdom to adjudicate conflicts between values and allow action to proceed. In the context of an ethics for the depressed,

I think Wong's concept of effective agency from Chapter Four may helpfully play this role.[46]

MFT is at its strongest in its defense of value pluralism and at its weakest in its defense of intuitionism. The philosophical discussion of value pluralism tends to understand values as concepts and value comparison as reflective, challenging monistic claims like, for example, that fairness and loyalty can only be intelligibly compared in terms of an ultimate value like well-being.[47] MFT understands pluralism as a psychological claim about intuitions, not concepts, and argues that the research into intuitive moral responses does not square with a monistic claim that moral intuitions are governed by a single value paradigm like care/harm or fairness/cheating. For example, pluralism better explains why "incidental disgust harshens moral judgments": disgust may be a response to a violation of sanctity even if that violation is harmless, suggesting that sanctity is intuitively valued on its own terms, not in terms of harm caused (Graham et al. 2013, 104). MFT provisionally proposes five value paradigms: care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and sanctity/degradation. These moral foundations are responses to adaptive challenges with characteristic emotions and relevant virtues with plausible evolutionary accounts for their original triggers and comparatively novel current triggers. For example, loyalty functions to form coalitions,

---

[46] Another candidate for understanding how values and emotions may be incorporated into an account of agency or freedom and responsibility is P. F. Strawson's (1962) account of reactive attitudes. But his account is tied to broader arguments about free will that I want to keep at arm's length. I emphasize the importance of distinguishing between reasons and causes to maintain a distinction between justification and motivation, because a depressed person's management of self-disregarding intuitions involves resisting the impulse to respond to a loss of motivation (from, say, a quasi-belief in one's own worthlessness and impossibility of success) with seeking justifications (like, say, with the futile attempt to refute a quasi-belief with competing beliefs). What matters to me is that depressed persons experience themselves as free, not that I can prove that they are. Such proofs matter to the skeptic. Someone afflicted with existential guilt, by contrast, does not need any convincing of their own responsibility.

[47] See Fitzpatrick 2022 for a more extensive discussion of analytic value pluralism.

responds to threats to the group, provokes pride and rage, and applies alike to nations and to sports teams or fandoms (67). The significance of these foundations, however, is contingent on the correctness of intuitionism. If agents are generally skilled at moderating their intuitions through deliberation and moral action is largely deliberative, then value paradigms might play little role in influencing moral action. By contrast, if successful deliberation is rare or limited in its capacity to moderate intuitions, then MFT's value paradigms could be a primary framework for understanding moral action.

The central evidence that MFT, and Haidt specifically, offer for intuitionism is the case of moral dumbfounding: persons who are refuted when attempting to justify their intuitive response to a perceived moral wrong may refuse to concede, doubling down on their intuition rather than modifying their view or seeking a new justification. Moral dumbfounding is observed when presenting hypotheticals that are likely to strike someone as intuitively immoral, such as a case of incest, but are structured to undermine an objector's likely justifications for their intuition: the experience is consensual and enduringly positive, precautions are taken and there are no negative consequences, etcetera. Haidt takes this as evidence that justification is generally rationalization: intuitions motivate moral responses and are less responsive than moral philosophers would like to moral reasoning. Valerie Tiberius (2015) rightly objects that moral theories dependent on rational moral reasoning, like Kantianism, distinguish between the theory and a decision procedure: all that matters is "that some moral judgments (the correct ones) are backed up by rational principles and that we *could*—if we needed to—use our rational capacities (such as universalization) to justify these judgments" (99). She also notes that Haidt's (2012) argument for intuitionism is tied to a normative argument for

the goodness of the intuitions that produce moral dumbfounding: for example, he thinks

that valuing sanctity, associated with disgust, "helps bind individuals into moral

communities" (268). This presumes that moral communities bound together by disgust of

what is considered profane, such as, say, flag-burning or homosexuality, are good.

Evaluating such communities in terms of their kindness or fairness might suggest

otherwise. Moreover, as Daniel Kelly and Nicolae Morar (2014) argue, disgust is

unreliable: it combines a strong sense of certainty with socially malleable objects, such

that one's disgust response to perceived poisons or pathogens is easily extended to out-

groups of people. If dumbfounding is the result of disgust responses, then it is less a

problem for moral reasoning than it is a problem for disgust as a justification for

beliefs.[48]

      Flanagan (2016) compares MFT to Mencius's "four sprouts" theory of moral

foundations and argues that while MFT better explains how virtues and vices may

emerge from the same foundation, Mencius's theory shows that MFT lacks a necessary

account of practical wisdom. Flanagan distinguishes MFT's pluralism from Haidt's

"hyperbolic" claims of intuitionism, though MFT is explicitly committed to intuitionism,

and frames it as a form of moral modularism (101). Like Mencius's sprouts, moral

---

[48] The heart of Haidt's conservatism, I think, is his belief in the goodness of disgust. In my own assessment, none of his examples of moral dumbfounding succeeded in dumbfounding me. For example, in response to his hypothetical regarding incest, I evaluated it as moderately wrong because of the moral hazard of a valuable source of trust between siblings being undermined by the loss of a platonic relationship. I judged that though such consequences were avoided by luck, the hazard remained. What matters is not if my judgment was well-reasoned but that I offered a justification not refuted by the example. But then I realized the reason for my lack of dumbfounding: I hadn't reacted with disgust. I had evaluated the example primarily through the paradigm of loyalty and betrayal. I thus think that MFT's distinguishing between loyalty and sanctity is useful, even if they are part of the same moral cluster, because only the latter seems to ride on disgust. I am inclined to see disgust as unfortunate evolutionary detritus, like the moral equivalent of the appendix. Maybe I am underrating its importance in maintaining group cohesion. But if it is necessary, insofar as kindness and fairness are the highest values, it is a necessary evil.

foundations are automatic and fast-acting; modify affect, judgment, and willing; and have "features of cognitive impenetrability" like being "hard to turn off" and stopped "only with considerable conscious effort/veto" (79). Flanagan notes three key differences between MFT's moral foundations and Mencius's four sprouts of benevolence (*ren*), righteousness (*yi*), propriety (*li*), and wisdom (*zhi*). First, fairness is missing from Mencius's account. Second, where MFT treats the foundations as equal, Mencius claims that benevolence and righteousness are higher than propriety, asserting a hierarchy of values. Third, wisdom, which Mencius claims is able to discover the hierarchy of values, is missing from MFT's account. The problem with the absence of wisdom from MFT, Flanagan states, is that "what we ought to do, how we ought to weight the outputs of the modules, is not determined by the modules themselves" (93). Intuitionism is thus necessary to explain how values can lead to action in the absence of any deliberative mediation: one's highest value just overrides the others.

Such a picture of ethical life seems descriptively inaccurate, in its absence of any substantive internal conflict and resolution between an agent's values (let alone sincere or self-transformative resolution of interpersonal disagreement), and normatively impoverished in its lack of any grounds for autonomy. Flanagan broadly affirms the evolutionary picture in which "some tendency worth cultivating … is judged to occur in something like a moral or protomoral guise," like how "disgust at sour milk, which is not moralized" might involve the same tendency as "disgust at incest, which is moralized" (108). But as Flanagan's emphasis on wisdom shows, the difference between the "moral" and the "protomoral" is the difference between reasons and causes. To think that someone being moved by disgust to a judgment that will not be revised even upon

256

dumbfounding is an example of moral reasoning is just to reject the foundations of ethics in some capacity of, if not free will, then intentional self-transformation. Even setting aside normative arguments, a description of moral psychology that lacks any reference to wisdom, as something that many people and most philosophers and sages at least think exists or act as though they do, is likely inaccurate. An ethics for the depressed, to maintain its needed distinction between motivation and justification, should maintain the distinction between the protomoral, that being values as pre-intentional, and the moral conceptualization of said values as norms.[49]

In an echo of the problem of the walled garden, Greene (2013) critiques Haidt as ignoring partisanship: American conservatives, for example, are not just defined by their greater valuing of the in-group/out-group values of loyalty, authority, and sanctity, but by the specific groups they identify as in-groups. He wryly notes that "American social conservatives are not especially respectful of Barack Hussein Obama, whose status as a native-born American, and thus a legitimate president, they have persistently challenged" (339). Being an authority figure is not sufficient to be worthy of deference, even if one values deference to authority, just as not all religions will be seen as sacred to those who value sanctity, as when birtherism is coupled with the claim that Obama is secretly a

---

[49] In a footnote, Flanagan mentions a conversation with Haidt he had at a bar in which "Haidt said that he thinks there might be a small band of people, a little tribe of philosophers for whom reason functions the way they say it can … But this function is very rare and next to impossible to inculcate in ordinary people" (301). As a founder of the Heterodox Academy, Haidt has been active since 2015 in advocating "viewpoint diversity" at universities, seeing academia as biased against conservatism. This cause is helpfully contextualized by awareness of his theoretical commitment to intuitionism and his personal commitment to elitism. The ideals of "free speech" and "open debate" are undercut by the intuitionist claim that moral reasoning is mere social strategy, unlikely to change beliefs or dispositions. If values are immutable traits and political affiliation is defined by values, then Haidt's aim of increasing the representation of conservatism in academia cannot be justified in terms of promoting constructive debate as an educational benefit. It could, however, be justified on the grounds that political affiliation should be considered a protected class, with conservatives seen as unable to change who they are and therefore not responsible for their views, rejection of which could amount to a kind of persecution. If Haidt thinks this, he should say so.

Muslim. The fact that one's awareness of one's group identity strongly filters the application of one's values strengthens the claim that an ethics for the depressed, as a partial ethics that begins deliberation with intuited values, depends on 1) an account of effective agency as practical identity (including group identity) and practical wisdom mediating values as pre-intentional and 2) on a politics for the depressed, understood as an account of how contingent value ranks defined in terms of a sense of stability can justify both self-transcendence and transcendence of communal values.

A full account of effective agency would explore the research on practical identity and distinguish its reflective dimension, or practical wisdom, from its intuitive dimension, or what determines which objects (including persons) seem relevant to a value paradigm. A full account of a politics for the depressed would answer the questions of 1) if or how felt value pluralism develops into political pluralism, 2) if or how the heart as the set of felt values relates to the person as the bearer of rights, and 3) if or how the guiding aim of a sense of stability extends to the pursuit of solidarity.[50] But a provisional account may be grounded in a return to Dewey's pragmatic meliorism, modified by the addition of the concepts of a qualitatively ambiguous situation and qualitatively opaque problem from Chapter 3.

To seek a sense of stability is to desire reliable or sustainable experiences of value. This resonates with Dewey's (1957) concept of growth as "adding fullness and distinctness of meaning" (259), which I take to also be the basis for a kind of gentle

---

[50] I believe that a good starting point for such an account might be Charles Mills's 2018 "Black Radical Kantianism," which reconceives respect for persons in the context of a racialized state defined by material and social inequality between normative and stigmatized groups. But an account of politics for the depressed remains a separate project.

perfectionism. While a qualitatively unified situation is defined by common felt values and is thus intuitively the "same situation" to those involved, a qualitatively ambiguous situation is defined by some of those involved perceiving a problem while others do not, either because 1) only some intuit the relevant values or 2) only some intuit those values as applying to the "problem," e.g., the persons or other matters over which the disagreement arises. Unlike with qualitatively unified situations, which may begin with the intuitions of the participants, the ideal resolution of a qualitatively opaque problem in terms of promoting growth involves demonstrating that there is a problem to those who do not intuit it. Insofar as normative argument is a fair and effective means of this necessary demonstration, deliberation should be understood equally in terms of dramatic rehearsal and normative argument. This revision both reconciles Dewey's politics with his epistemology and reconciles classical Deweyan pragmatism with the Rortyian neo-pragmatic account of normativity as a matter of the exchange of reasons, not the force of causes.

4. Ethics of Engagement as Pragmatic Meliorism

A Deweyan or Jamesian pragmatic ethics is suited to being an ethics for the depressed in restricting itself to hypothetical imperatives, treating the moral and prudential as continuous, and emphasizing pre-intentional states such as habit. In fact, Dewey's account of prizing in his *Theory of Valuation*, relying on his previously discussed account of habit, is a theory of values as pre-intentional, and his account of deliberation as dramatic rehearsal resembles a non-transcendental phenomenological method. But there are two challenges in incorporating a pragmatic ethics into an ethics

259

for the depressed. First, the primary focus of pragmatic evaluation is conduct, broadly construed, and depressed persons struggle to act. If pragmatic methods are defined primarily by the assessment of action or behavior, the depressed, in seeming inert, may be inscrutable. Second, qualitatively vivid intuitions or feelings can be misleading to depressed persons and thereby cause the problems of demoralization and hypermoralization. To avoid a downward spiral into rumination, depressed persons cultivate a vigilance regarding their own impulses and their suitability for initiating deliberation, amounting to a suspicion of one's felt sense of a problematic situation. For the depressed, to begin deliberation with a feeling, insofar as that feeling is more likely to be unyielding anxiety or dread, is less likely to lead to growth.

I have sought to meet these two challenges in two ways. First, I have understood demoralization and hypermoralization as arising from pre-intentional states, including habits, such that a Deweyan framework better illustrates what goes on in depressive loss of motivation than a traditional account of action guided primarily by intentional states. This situates the notion of a "habit of feeling" as potentially continuous with a broad pragmatic notion of conduct and as compatible with Dewey's theory of habit. Second, I have modified Dewey's account of deliberation to distinguish between qualitatively unified and qualitatively ambiguous problematic situations, thereby resolving the tension between his epistemology of affectivity and his egalitarian politics. As I argued, without this distinction, persons with different qualitative experiences of the unity of a situation are necessarily in different situations and thus cannot resolve the same problem. I take Dewey to be elevating the importance of feelings to successful deliberation rather than to be diminishing the importance of rationality. Even so, clarifying the category of

260

qualitatively ambiguous situations and their qualitatively opaque problems situates the importance of normative argument, in contrast to dramatic rehearsal, as a fair and effective response to such situations. In contexts where fair debate is ineffective, as when one is viewed by one's opponents as irrational or ineligible to participate in debate, non-argumentative responses like protest become understandable and justifiable.

For Dewey, as for depressed persons, the experience of conflict is inevitable. "All action is an invasion of the future, of the unknown," he writes: "Conflict and uncertainty are ultimate traits" (1957, 12). For depressed persons, the sense of precarity remains even as a future seems impossible. It is thus possible for both Deweyans and depressed persons to fail to recognize that sometimes conflict must be actively uncovered or intentionally evoked if growth is to be possible. The category of qualitatively opaque problems encompasses such cases.

I am inspired by the non-ideal theoretical interpreters of Dewey's ethics who I take to emphasize meliorism over self-transformation or an ideal normative procedure, a group that includes Colin Koopman, Mark Johnson, and Elizabeth Anderson. Non-ideal pragmatic ethics tends to focus on implicit normative content and structure, as in skillful action or in metaphorical language, as what is made explicit in norms as principles or propositions. Ideal principles are derivative of the primary matter of ethics: the modification of implicit norms through, say, change in habits. To know in what way to change habits, non-ideal pragmatic ethics may propose, like Dewey, an ur-value such as growth or conscientiousness as a normative analogue to Darwinism. Rather than attempting to prove that persons should value this ur-value, non-ideal pragmatic ethics turns to descriptions of what people actually do value, as shown through implicit

normativity, and presents the ur-value as an explanatory category: in enhancing

experiences of value overall and in the long run, it represents what people are already

aiming for in more or less effective ways. Whether this descriptive ethics involves a

naturalistic turn to science or a genealogical turn to archives, and whether it explicitly

posits an ur-value or not, the emphasis is on what persons already value as the basis for

understanding how to transform values, individually and societally. This is pragmatic

meliorism, which is distinct from readings of Dewey that focus on self-transformation as

such, not criteria that might determine what counts as valuable self-transformation, or

that supplement Dewey with an ideal normative procedure.[51] Though it is an example of

a self-transformative approach to pragmatic ethics, I am also influenced by Sarin

Marchetti's exhortative ethics and its relation to Nora Hämäläinen's account of self-help.

---

[51] For reasons of space, I must relegate my readings of the self-transformative approaches of Steven Fesmire, Gregory Fernando Pappas, and Todd Lekan, and the ideal normative approach of Philip Kitcher, to this footnote. Fesmire (2003) presents dramatic rehearsal, "*crystallizing possibilities and transforming them into directive hypotheses*," as identical with Deweyan deliberation (70). I concur with J. E. Tiles's (2004) observation that this disregards the subsequent reflective phase of assessment. Pappas (2008) is strongly committed to the situation as being essentially qualitative and takes it as the basis of a radical particularism, "an ethics based on the capacity of each situation to guide and rectify its own problems and challenges" (301). Tiles (2009) notes that though Pappas tries to distinguish his "primary experience" from "pure experience" because primary experience has "conceptual and cultural baggage," Pappas still claims that primary experience is not "inside a language, a culture or a socioeconomic system" (2008, 22), which seems to be a distinction without a difference. Fesmire and Pappas focus on pragmatism's particularism, its responsiveness to unique situations, and comparatively ignore Dewey's naturalism, but the latter serves as the basis of the general claims that guide how one should respond to situations. Lekan (2006) redefines Deweyan deliberation to attend to "the distinction between explanation and justification" by distinguishing "decision problems," solved with "antecedently justified procedures," and "determination problems," the response to which is "to modify habits in a novel way" (258). But he does not apply this distinction in his book on Deweyan ethics, where he claims that "the terms *norms* and *habits* may be used interchangeably" (2003, 61): explicit normativity, and thus the realm of justification, is disregarded in the focus on implicit normativity. In contrast, Philip Kitcher (2011) presents a naturalistic account of ethics as a response to "altruism failure" (98), not as a matter of moral facts, and constructs an ideal standard for assessing ethical progress that includes "replicating the course of an ideal deliberation under conditions of mutual engagement in which all members of the human population participate" (343). Grant that altruism is the dominant value shaping group dynamics, setting aside the distinct evolutionary roles of fairness, loyalty, and sanctity. Even then, as William FitzPatrick (2012) points out, the "historical fact" of altruism is not a reason why someone who rejects altruism should accept it (173). Alex Sager (2014) also notes a tension between Kitcher's development of an ideal procedure and his evolutionary account, which lacks examples of a society developing in accord with an ideal procedure. I regret that I cannot say more here.

Given my concern with a depressed interlocutor, I situate argument as having a legitimate

exhortative function rather than, say, opposing philosophy as idealized argument with

persuasion or sophistry. But I worry that self-transformative pragmatic ethics puts both

too much stock in the transformative power of theory and too little in the capacity of

pragmatic ethics to justify normative reasons as distinct from inducing re-habituation or

instilling confidence.[52]

Like Marchetti (n.d.), I hope to legitimize descriptive and hortatory ethics

alongside prescriptive ethics as properly philosophical, if only to reveal perspectives

overlooked in prescriptive ethics. Hämäläinen's (2015) review of the sociology on self-

help, which he cites, finds illuminating parallels between self-help and moral philosophy:

> [self-help] is (to a large extent rightly) accused of being overly concerned with
> individual persons, seeking individual solutions, and disregarding the varieties of
> ways that people's lives are interconnected and limited by circumstances. Its
> conception of the human person is generally atomistic, thin, "unencumbered." It is
> in fact, to a large extent, the same thin idea of the human person that Iris
> Murdoch, in the 1950's and 60's, found problematic among her philosophical
> contemporaries … Self-help and modern moral philosophy are in important
> respects products of one and the same culture: modern, western, predominantly
> secular. A thin liberal/modern conception of the human person is one of the things
> that they in central respects share. In philosophy it might work as a respectable
> abstraction, designed to emphasize the moral and existential freedom and dignity
> of individuals, beyond contingent practical limitations, traditions, and cultures. In
> the criticism of the self-help debate this modern abstraction shows its less
> flattering face: it is not so much an ideal abstraction as it is a falsification (perhaps
> unwittingly) imposed on people to make them bear individual responsibility for
> things that often are beyond their control: illness, unemployment, an increasing

---

[52] Again, for reasons of space, I restrict this section to a discussion of Dewey's own theory of value and not
the contemporary pragmatists who have most influenced me, but I would be remiss in failing to
acknowledge them. In my turn toward the pre-intentional as the center of value ethics, I am inspired by
Koopman's (2017) Jamesian concept of "willful rehabituation" as "an experimental attitude that affirms
possibility" (494) and attention to the implicit normativity of "embodied skill" like "successful dance
performance" (2011, 75). My emphasis on effective agency and practical wisdom is guided by Johnson's
(2014) discussion of conscientiousness as "the mental and emotional flexibility to imagine new solutions"
(216) and vision of an ethics in which "growth in depth, richness, and scope of meaning" is "your only
guide" (200). In future work on a politics for the depressed, I hope to draw on Anderson's (2009)
discussion of relational egalitarianism from a "non-ideal standpoint methodology" (138).

demand for "social competence," and increasingly fractured social relations. (15-16)

The problem with self-help is not that it is vulgar but that it is too often a grift. Like a fad diet, it overpromises, underdelivers, and obscures its limitations or conditions of success so that when it fails, the agent is disposed to blame themselves rather than the method. What should distinguish ethical theory is not that it is only evaluated as prescription and neither as description nor exhortation, but that it underpromises, overdelivers, and exposes its limitations for criticism.

This anti-elitism and primacy of the practical would have appealed to Dewey. It is to his theory of valuation I turn to conclude the sense in which an ethics for the depressed is a form of pragmatic meliorism: as an ethics of engagement for persons predisposed to altruism and alienation. The challenges that Dewey faces in linking his epistemology and his politics foretell the challenges of politics for the depressed and of generalizing an ethics of engagement.

Dewey's (2008) distinction between prizing and appraisal as stages of valuation introduces an understanding of normative reasons as "ends-in-view" that, in "examination of their respective conditions and consequences," may be gradually improved (245). Ends-in-view are distinct from ends-in-themselves, which Dewey rejects as an "astonishing" notion, in being contingent responses to an agent's environment and a means to future ends-in-view. An end-in-view is "*warranted*" (213) as a proposition or "appraised or valued as *good* or *bad*" based on its capacity to resolve "some lack or conflict" found in a situation. The distinction between what one happens to value, or prize, and what one should value, as a product of appraisal, is for Dewey an empirical matter, not requiring some separate category of the normative. "The contrast referred to,"

he writes, "is simply that between the object of a desire as it first presents itself (because of the existing mechanism of impulses and habits) and the object of desire which emerges as a revision of the first appearing impulse, after the latter is critically judged in reference to the conditions which will decide the actual result" (219). In this account of value, "the head and the heart work together," which is to say that "prizing and appraising unite in direction of action" (249). This is the previously cited Deweyan view of deliberation as "rational in the degree in which forethought flexibly remakes old aims and habits, institutes perception and love of new ends and acts" (1957, 186). By its lights, "morals means growth of conduct in meaning" (259): it is enough "to stimulate us to remedial action, to endeavor in order to convert strife into harmony, monotony into a variegated scene, and limitation into expansion" (260). The ur-value of growth, like Scheler's concept of love, is not itself a value so much as it is what drives the increase in value, which for Dewey is through the resolution of conflict, the introduction of novelty, and the broadening of horizons. But unlike Schelerian love, it is not just movement upward but outward, toward a wider set of objects that may be found valuable.

Dewey explicitly rejects the need for any account of normative authority separate from a descriptive account of what persons value and by what means they may best realize those values:

> Still the question recurs: What authority have standards and ideas which have originated in this way? What claim have they upon us? … The authority is that of life. Why employ language, cultivate literature, acquire and develop science, sustain industry, and submit to the refinements of art? To ask these questions is equivalent to asking: Why live? And the only answer is that if one is going to live one must live a life of which these things form the substance. The only question having sense which can be asked is *how* we are going to use and be used by these things, not whether we are going to use them. (75)

Though Dewey's vision of philosophy is as continuous with science and Williams's vision of philosophy is, like Rorty's, normative and argumentative, the two come to a similar conclusion: a philosopher cannot prove that life is worth living to one who doubts this. It is because ethics cannot coherently or sustainably begin with an idealized principle or procedure unmoored from psychology and biology, or our best understandings of them, that ethics must instead begin with an agent who already has values and projects, not one who must be argued into having them. I agree. This claim is compatible with an ethics for the depressed because depressed persons do not lack values, projects, or desires: rather, these are undercut by conflicting quasi-beliefs, habits, or dispositions. Dewey's view of "existence as precarious and as stable" (1958, 40), of conflict and threat as facts of life, also resonates with the depressed sensitivity to the sense of stability.

But two problems arise. First, some experiences of precarity cannot be relied on to initiate deliberation: existential guilt spirals into rumination instead. Second, sometimes persons must be persuaded that precarity exists precisely where they do not experience it. A fair means of doing this is by giving reasons, which depends upon a concept of reasons as distinct from causes. For a pragmatic ethics to be an ethics for the depressed, then, it must be supplemented by a phenomenology of value and intention formation, to better understand what occurs in depressive habits prior to observable action. As pragmatism becomes phenomenological, so phenomenology becomes pragmatic, renouncing a transcendental method in favor of an introspective kind of immanent critique. And in modifying's Dewey's understandings of the situation to include the qualitatively ambiguous and of deliberation to emphasize normative argument

as a response to qualitatively opaque problems, hopefully thereby reconciling classical and neo-pragmatic concerns, normative argument is situated in an exhortative context. When a depressed person breaks a spiral of demoralization and hypermoralization, the subsequent experience of themselves as rational, as capable of forming intentions and giving reasons, is truly gratifying and empowering. It is in search of this experience that depressed persons, even when subject to compulsions unresponsive to argument, meaningfully seek reasons. They seek their autonomous selves.

Contingent value ranks, which rest on the claim that the less partial values of kindness and fairness are overall and in the long run more conducive to a sense of stability, may be generalizable beyond depressed persons to the extent that a sense of stability is generally sought. But depressed persons respond to a sense of precarity that is hard to suppress. The intransigence of habit, especially when reinforced by one's community or ideology, is otherwise effective at suppressing one's awareness of precarity and thus of problems. This derailing of inquiry is, as I discussed in Chapter Three, the challenge that Dewey attempts to confront in his egalitarian response to the elitist rejection of democracy in *The Public and its Problems*. I take the possibility of ideology suppressing the experience of conflict to represent the tragic risk of, in the words of Cornel West (1989), "wholesale regression" instead of growth (57). A fuller defense of contingent value ranks as a normative ethics depends on a political philosophy of engagement, including a lengthier discussion of effective agency as pertains to political or group identity.

In brief, to resolve the problem that "lower" values may be more productive of a sense of stability than "higher" ones, an ethics of engagement must account for, as in

Flanagan's critique of Williams, the possible goodness of engagement's negation: alienation. For those neither 1) predisposed to especially value the less partial values of kindness and fairness nor 2) predisposed to alienation, felt values may persistently incline toward in-group partiality. As is, the theory of contingent value ranks only affirms this failure of transcendence. An ethics of engagement, then, might paradoxically require the inducement of alienation in others from their communities to prepare the way for a form of engagement defined by higher values: solidarity. But justifying this requirement is akin to considering the application of an ethics for the depressed to the egoist or the skeptic, the more traditional interlocutors with whom I declined to begin this theory. A generalized ethics of engagement requires a solution to the problem of the walled garden that continues Dewey's political project. For depressed persons, who begin with tendencies toward altruism and alienation, there is instead a different and distinct answer: the "call of the other."

CHAPTER VI: THE CALL OF THE OTHER

I began all this with a bet: I would develop a recognizable ethics from existential guilt. I have taken this guilt to be the defining feature of a depressed person: the compulsive and pervasive sense that one is ethically lowered or unworthy for no specific reason. There are causes of existential guilt, like brain chemistry or trauma, but the causes of existential guilt are not justifying reasons for existential guilt. However, the one experiencing existential guilt is compelled to think or feel that there must be a reason to justify the experience, which is to say that they come to believe that they must be guilty of something (or everything). Unlike standard or intentional guilt, existential guilt is not a response to a specific wrong one believes oneself to have committed. Thus, with no wrong to address, it seems to lack the prospect of forgiveness.

As compulsive, existential guilt is an unreliable indicator of one's own ethical status. It involves a set of deceptive intuitions about 1) altruism, like that one must give and sacrifice and seek some vague redemption because of one's special unworthiness, and about 2) one's own alienation, like that this requirement of sacrificial giving is unique to oneself, so one must hold oneself to some vague higher standard, intensifying one's sense of isolation from others and of being especially unworthy. Since existential guilt is a deceptive compulsion, it is resistant to refutation by reasons. It is thus reasonable to think that philosophical argument or ethical theory will not be helpful in alleviating existential guilt, or more bluntly, that ethical theory and depressed persons have little to gain from one another. I have aimed to show the opposite.

Ethical theory, understood as purely or primarily prescriptive, is directed and limited in its prescriptive concerns by its descriptive and exhortative dimensions. Because

metaethics and moral psychology have mostly addressed the already autonomous agent, like the egoist or the skeptic, they have largely overlooked mental states that the already autonomous agent takes for granted but that may be modified in "aspiringly autonomous" agents like the depressed. These states include the pre-intentional conditions of motivation and deliberation that lead to demoralization and hypermoralization. In turn, ethical theories that aim to eliminate moral luck and justify the objectivity of moral obligation, meaning an absolute measure of ethical merit and blame, may ignore their dependence on pre-intentional conditions like habits. This leads to unwittingly promising an impartial measure of moral status to the depressed, who long for it as a counter to their compulsive self-devaluation, and then not delivering. Depressed persons, though they cannot be argued out of being depressed, can benefit from the descriptive and exhortative dimensions of ethical prescription delivered as prescriptive. Or, more precisely, an ethics for the depressed is meant to be exhortative in every dimension: as descriptive, as prescriptive, and as lived through or learned. I will now describe these three dimensions in greater detail.

By the descriptive dimension of ethics, I mean what Paul Biegler (2010) calls the material facts of one's condition. For an ethics for the depressed, these are the accounts of demoralization and hypermoralization. Demoralization should be understood not as the loss of desires or goals but as the presence of a defeating pre-intentional state, like a quasi-belief in the impossibility of success or a habit of feeling overwhelming dread or anxiety, that interrupts the translation of motivational states like desires into intention formation or action. Depressed persons are not listless because they stop caring but because they experience a powerlessness to act on their concerns. Hypermoralization

should be understood as a practical difficulty in maintaining a categorical distinction between moral obligations and prudential interests that is linked to self-alienation: what one wants to do becomes perceived as what one "must" do, prior to any justification for thinking so, and deliberation becomes spiraling or rumination, in part because one loses the sense of one's own distinctness as an individual with worthy projects and purposes. These explanatory accounts do not themselves justify any normative claim that, say, demoralization or hypermoralization are good or bad. That gap between descriptive and normative is, in fact, what benefits depressed persons as compulsive moralizers. These accounts may help a depressed person to reconceive themselves in a more descriptively neutral, less normatively evaluative way: as subject to compulsions rather than sinful or cursed or foolish or fallen. I thus hope not just that these accounts are accurate but that they are encouraging.

The prescriptions of an ethics for the depressed are responses to existential guilt and responsive to depressed persons. In being responses to existential guilt, the prescriptions are constrained by a sense of what is possible for depressed persons as agents, guided by the above accounts of demoralization and hypermoralization. Thus, a value ethics of engagement begins with a relatively reliable source of motivation for depressed persons, the intuitive response to a problematic situation prior to any deliberation or rumination, and does not require a decision procedure that draws a distinction between moral obligations and prudential interests. In being responsive to depressed persons, the ethics is motivated by some of the same intuitions it is meant to help resist. The argument for contingent value ranking, that the less partial values of compassion and fairness should be seen as higher than the more partial values like loyalty

271

and sanctity, is justified by a conceptual claim about how the value-concepts are used and an empirical psychological claim about how the values-as-felt produce a desirable sense of stability. But there are two separate reasons to hope the argument succeeds. The first is that an ethic that elevates less partial values, in tending toward self-transcendence and the transcendence of communal values, is more recognizably an ethic and not a mere affirmation of priors. The second is that depressed persons, because of existential guilt, tend to be excessively strict on themselves. One goal of an ethics for the depressed, insofar as that strictness inhibits planning and action, should be to ease that guilt. I think that the more an ethical theory can justify the goodness of self-transcendence, communal transcendence, and less partial values, the more plausible of a substitute for strictness it is likely to seem, even to persons under self-disregarding compulsions.

In grounding an ethical theory on the deceptive intuitions of existential guilt, that guilt motivates an argument for its negation: self-trust. It is because depressed persons are inclined toward altruism and alienation, because they are existentially guilty, that they are inclined to aim toward transcendence, to doubt themselves and question their assumptions. Doubt and questioning only become a problem for depressed persons when the absence of a satisfying answer makes it impossible to go on. If the risk of acting wrongly seems to outweigh any possible reward, then a depressed person does not act. Becoming willing and able to act means changing one's habits and learning how best to respond to one's compulsive thoughts and feelings. But after these habits are changed and responses are learned, the doubt and questioning remain. Depressed persons who successfully cope with depression have not lost their tendencies toward altruism or

alienation. And, for an ethics for the depressed, it is those very tendencies that serve as the justification for a partial ethics, a reliable ethics, an ethics that goes easier.

There is no such thing as redemption from existential guilt because it does not issue from any wrong committed. It is not countered by arguments. It does not respond to reasons. At the practical level, day to day, moment to moment, existential guilt is countered by indirect responses: it is softened, eased, unwound, by gentle questions, deep breaths, confiding in others. Yet this theory does offer a redemption of a different sort. In framing existential guilt as the basis of an ethical life, an ethics for the depressed redeems that guilt. Existential guilt is itself not sinful or cursed or foolish or fallen. It is good. It makes people good. It just needs understanding.

Of course, its urgings are unreliable, as prone to self-disregard as regard for others. But it is an acceptable starting point for an ethics of engagement. Demonstrating that point is the greatest exhortation that I think this project offers. It shows that a depressed person can be good.

To philosophers, I hope it becomes clearer that there is a middle ground between those who are responsive to reasons – to do good, to live well, to go on – and those who are not: those who desire to be responsive to reasons. The ethics of care and feminist philosophy have drawn attention to the forms of dependence that underlie autonomy and have presented autonomy as transitional, attained in the passage from child to adult. I am far from the first to notice that when philosophy speaks only to the already autonomous, it speaks from, to, and for privilege: from and to those who were cared for and who may have passed off care work to others, for those who may be unconcerned with edge cases of agency like children. A depressed person, understood as defined by existential guilt,

273

might be thought of as a different sort of edge case of agency: one who remembers being rational, who is in fact rational and still reasoning, but who is actively undercut from within throughout the reasoning process. Depressed persons are already acquainted with the object of their desire. To legitimize them as an interlocutor is, I think, one more step toward a fuller reckoning with reason-giving and the conditions of its possibility.

An ethics for the depressed should help the agent to become an agent. This is possible. It may even be necessary. But arguments themselves are not enough, because arguments are best for those who are already agents. Like education itself, the cultivation of agency is a challenge rife with opportunities for manipulation and abuse. There are always assumptions not argued for or habits inculcated before they are justified, without presenting the recipient with the opportunity to evaluate these assumptions or habits in advance. But reason-giving need not require an ideal context if it aims toward fairness and compassion, and thus toward the creation of the very context from which it benefits as reasoning. What is necessary above all else is trust: trust that failure is tolerable, that provisional success remains success, that there will be time to improve, that there is a future, that we may try again.

There are good reasons to doubt these claims and there are circumstances in which they are false. It is unfair to require trust in an untrustworthy world. The world must be made more trustworthy, more stable, so that it is more reasonable to trust in others and in one's own future. But if we are to make that world, or maintain that world, I think some trust must come first.

In this final chapter, I aim to understand this trust, what Jay Bernstein, following Annette Baier among others, calls "trust in the world." Recalling Michael Cholbi's claim

with which I began, that depressed persons may be especially altruistic, I position what he calls "altruism" as existential guilt manifesting as "the call of the other," which is a combination of inclinations toward sympathy and alienation, and neither just sympathy nor some immediate contact or communication with another subject. I frame existential guilt and the call of the other as means by which an agent's embodied vulnerability, as the structures of dependence that serve as the foundation of autonomy, may be revealed to them. Trust in the world is initially a lack of awareness of one's own vulnerability, that is, one's capacity to have one's autonomy – not just one's actions, but one's very will – undermined by others. The momentary forgetting or disregarding of one's vulnerability is common to, and might even seem a necessity for, engaged and skillful action. But an agent aware of the vulnerability of themselves and others may strive to create a world that is more trustworthy, more preserving of trust. Such action is not premised on the continued forgetting or disregarding of vulnerability but instead on trust.

I use the term "alienation" loosely to mean a state that tends to evoke feelings of loneliness or isolation. I noted in Chapter 5 that to generalize an ethics for the depressed into a value ethics of engagement would require a lengthier discussion of practical identity and the relation between group identity and value paradigms in structuring intuitions of value. Rahel Jaeggi (2014) defines the concept of alienation in a way that is helpful for this discussion: it is a failure to appropriate oneself or one's world, that is, to be able to feel oneself to be making oneself or one's world one's own. Jaeggi's appropriative model of the self contrasts with those of Charles Taylor, Bernard Williams, and Alasdair MacIntyre, and the concept of alienation as relevant here would be further clarified by a comparative reading of these philosophers. But it is illuminating to define

275

alienation in terms of a loss of individuation. In the context of an ethics for the depressed, the sympathy that arises in the call of the other may seem like an effective response to alienation not because it necessarily eases feelings of loneliness or isolation, but because the call of the other seems to endow sympathy with the capacity to individuate the depressed person. Through altruism, the depressed person's individual life seems to matter again.

To discuss the call of the other in the context of vulnerability and trust, I distinguish three claims that I associate with the call of the other. First, "the other is real when the self is not": a depressed person may be roused from the depths of profound alienation from others and themselves by a strong and desperate sympathetic impulse. Second, the call of the other thus functions for a depressed person as the essence of responsiveness: it is the experience of a rupture in what seems valueless and a return to what seems valuable. This explains its significance to the ontology or ethics of phenomenologists like Heidegger and Levinas. Third, the call of the other is nevertheless structured by value paradigms: the "other" to which one attends, the objects and structure of one's attention, is at least partly constrained by one's practical identity and pre-intentional values. Thus, the call of the other is not in fact responsiveness as such, insofar as this would have to precede one's contingent values. It is also not a justification for action but a motive of action, not a reason but a cause. One's conceptualized values and associated norms are one's reasons. A philosophical account of ideal normativity thus needs not include or depend upon an account of the call of the other.

Having argued that the call of the other is not immediate contact with another subject, I conclude with a reflection on Steven Crowell's treatment of Levinas and the

concept of the "face" as responsiveness as such to alterity, akin to the "call to conscience." The call of the other both 1) reveals the vulnerability that structures autonomy and 2) evokes compassionate action in its sympathetic quality and transcendence of self and community in its alienating quality. As such, it can function for a depressed person as something like "responsiveness as such," the essence or ground of all reasons for action. But as I have said, it is ultimately a cause of action and reason-giving, not a reason for action or reason-giving: it provides ultimate motivation, not ultimate justification. In Chapter 5, I noted that Crowell's concern with responsiveness as such resonates with Derrida's side of his debate with Rorty over the necessity of a transcendental or quasi-transcendental account of normativity. In the end, I favor Rorty over Derrida and so, probably, Kant over Hegel. If the work of ethics lies in making choices, then to give an account of the ways in which one must necessarily be ethical is not to do the work of ethics. And an essential aim of an ethics for the depressed, in particular, is to realize and sustain the experience of oneself as choosing, the experience of agency itself. As to if that experience reflects a real free will: to a pragmatist like me, the sustaining of an experience may be as close as one gets to knowing what is real. Or, at least, it is close enough to warrant a bet.[53]

---

[53] As Colin Koopman has explained to me in correspondence, both Rorty and Derrida would endorse the Hegelian point that reason must come from somewhere, whether history or care work, and it will be constrained by its conditions. However, Rorty would say that these conditions of reasoning cannot be known in advance of reasoning. Thus, to begin reasoning, it must be possible to reason without (or prior to) knowledge of the conditions of reason. This point bears on my discussion of a politics for the depressed, also in Chapter 5. To the extent that an ethics for the depressed is relevant to a politics, it is because of two previously mentioned ideas. First, although the context and motives of autonomous reason-giving are not internal to autonomous reason-giving because they are causes, not themselves reasons, that context should somehow contextualize and motivate the *account* of autonomous reason-giving. Second, a primary condition of the possibility of autonomy is trust. To take a cue from Jean-Jacques Rousseau's *Discourse on Inequality*, the central question of politics is that of legitimacy, which is to say, if there is any meaningful distinction of legitimacy between rule by coercion and rule by consent. The associated question, if agents are just brutely compelled to pursue their personal interests or if they may choose against their personal

In phenomenological ontology, the call of the other takes on a grandiose role.[54] I

agree with Crowell's assessment that, in the forms of Heidegger's "call to conscience"

and Levinas's "face," the call of the other is framed as responsiveness as such. I hope to

show that, first, the call of the other is psychologically simple: it is just sympathy and

alienation, with sympathy appearing as the solution to the problem posed by alienation.

However, second, because one is most attentive to it in a state of anxious breakdown, the

call of the other can both appear and function as a kind of ethical last stand or last resort.

Thus, third, despite the call of the other's unglamorous simplicity, to position the call of

the other as ontologically or ethically primordial is not exactly mistaken. It is a cause that

is very likely to be taken as a reason, not a reason sent from on high, so to call it

"responsiveness as such" may be misleading. Still, it does not just appear as sent from on

high: it functions in that way, as a reason that can still seem good enough for going on

when few or no others do. In this sense, phenomenology is right to call attention to the

phenomenon of the call of the other. Where it goes astray is in a refusal to look elsewhere

than the transcendental conditions of consciousness for an account of normativity.

I began the previous chapter with an anecdote about my habits of cleaning,

intended to exemplify a value ethics of engagement in practice. I meant for it to show

interests toward impersonal goods, resembles the central question of an ethics for the depressed, that being
if arguments or reasons have any relevance to the depressive loss of motivation. An ethics for the depressed
might thus answer the political question in a similar way: if one begins by trusting that an agent can reason
toward impersonal goods, one might be able to create the conditions under which it is true that an agent can
reason toward impersonal goods. This parallels my move to elevate normative argument as more important
than in Dewey's epistemology, even though it remains contingent on non-argumentative conditions like
habit formation, because it is politically fairer than the alternatives. If the choice to engage in debate may
establish a fairness between the participants that might not preexist the debate, then trust may serve as the
foundation of new respect. As we trust in each other despite our constraints, so we trust in reason despite its
constraints. But all this remains provisional and abstract.

[54] My 2014 master's thesis, "Continuity and Receptivity: The Systematic Intersections and Tensions
between Phenomenology and Pragmatism," focused on phenomenological ontology.

how a depressed person can pursue a good that is experienced as personal or respond to a problematic situation with a pervasive unifying quality. But I have not yet similarly shown how a depressed person can pursue a good that is experienced as impersonal or respond to a problematic situation that is qualitatively ambiguous.[55] To frame my discussion of the call of the other, then, I will begin with an anecdote about my best and most recent efforts to act altruistically in a context that might seem to lack clear or direct motivation, obligation, or gratification: my giving blood.

1. Giving Blood as an Impersonal Good

I have a fear of needles. Rather, I have a fear that I associate with needles: the fear of my own blood pumping. Even taking my own pulse is unsettling. Blood pressure tests make my skin crawl as the cuff tightens and I am made aware of my veins. But I do not seem to be especially afraid of the sight of blood. I think, therefore, that I am shocked by the reminder that I am made of parts. Here I am, struggling to manage an unruly world, but at least able to count on the wholeness of my body. And then it turns out that my body, all along, has been made of pieces that can fall apart. It strikes me as a second world for which I am responsible but about which I have been unvigilant. What I call fear is more a combination of cosmic horror and resentment. I thought that my body was as

---

[55] I discuss the distinction between problematic situations with or without a pervasive unifying quality in Chapter 3. I think that the capacity to respond to a problematic situation without a pervasive unifying quality is less involves one's practical wisdom or conscientiousness, in Mark Johnson's sense of the term as referenced in Chapter 5, and more involves reflection upon one's knowledge of context, which one is motivated to acquire by the less partial values. I limit an ethics for the depressed to a discussion of personal goods, as in cleaning, and impersonal goods that are personally motivated, as in giving blood. For a politics for the depressed, I would develop an example of an impersonal good in which one's personal stake is ambiguous, like the example of "B," the third party observing the possibly offensive joke, in Chapter 3. In such a case, one must be guided primarily by reflection, not by intuition.

whole on the inside as it seems from the outside. Frankly, I feel betrayed. Surely, I have enough concerns already.

Soon, of course, I feel ashamed. On reflection, I know that I can experience my body as whole because of my luck at not, say, suffering from chronic pain or other serious illness, or having the objectification of my body be a social or legal norm that I experience as constraining my freedom. My unearned sense of betrayal is typical of the fortunate person who is both unaware of their own luck and incentivized to remain so unaware. I hold such people in contempt when that feeling guides their actions toward cruelty and unfairness. While I am not aware of this feeling guiding my own actions in this way, I would be more assured on that point if I were to succeed in explicitly acting counter to that feeling. Anyway, frankly, I dislike this fear. On top of potentially reflecting poorly on me, as at least embarrassing if not disgraceful, it is unpleasant and seems beyond my control. I would be especially satisfied to overcome it.

So, I decided to overcome it through exposure: I would give blood as often as possible, which is about once every two months. The first time, the nurses took one look at my face and went straight to crisis mitigation mode, tilting back my chair, giving me ice packs and beverages, and quizzing me what felt like multiple times a minute, not that I was able to properly judge the passage of time in my state then. The second time, I had practiced deep breathing and taking my pulse to desensitize myself to my fear. Then, when I told the trainee nurse that I was OK, they believed me. That may just have been because the nurse was a trainee, but I called it a win.

What troubled me most was the form I had to sign, which I had once known about but had forgotten was required, indicating if I had recently had sexual intercourse with

another man or a sex worker. My sense of fairness rebelled. The question is as senseless as it is cruel. These persons, for obvious historical and material reasons, would be more likely than most to be conscious of the risk of infections transmissible by blood. The prompt is a manifestation of stigma, nothing more, a brute repetition of unjustified contempt, meant to mark its targets as essentially diseased and contagiously profane. I am bisexual, and checking that box felt to me like a disavowal, a betrayal of my people. But I judged that nothing would have been gained by my refusal to sign. I was there to help strangers and to better myself, not to champion my community.

Notice how I have done what I have done. In giving blood, I am personally motivated by the desire to experience myself more consistently as an agent, as well as related goals like the avoidance of shame. These motives are not necessarily self-transcending or transcending of communal values. The same motives could drive me toward conformity with a like-minded community that dogmatically affirms my already held values and that would shame me for deviating from them. But, with an awareness that my motives might guide me to a failure of self-transcendence and a desire to put them and myself to the test, I was able to direct my motives toward a higher value of altruism toward anonymous strangers to whom I will remain anonymous. Moreover, I was faced with a situation that I experienced as a conflict between in-group values and out-group values: the question of if I should act in a way that intuitively strikes me as traitorous and corrupt for the sake of compassion toward strangers whom I will never know and who will never know me. I will explain the significance of this with a hypothetical.

Imagine that I go on social media and post an account of this experience, and I encounter an advocate against the stigmatizing laws who says that I am in the wrong. This advocate is organizing an act of protest through mass refusal: no one who participates in their protest will donate blood unless the laws are changed. The prospect of group action rather than individual refusal complicates my scenario, not just because it is more likely to be effective, but because this means that out-group values are more clearly at stake. The cruelty and unfairness of intentionally withholding blood donations to those in need might conceivably be mitigated, or even outweighed, by the repeal of a cruel and unfair regulation. For an ethics for the depressed, this hypothetical advocate's plan is not absurd. It merits consideration. My response would be to appeal to a better way of achieving the same goal, one that is itself less cruel and unfair, and to look for an example of a form of protest that does not inflict the brunt of its impact upon those who are suffering most and least responsible. Of course, real advocates against these laws are not so clumsy as my hypothetical one. But social media takes all kinds, especially the strident and short-sighted, so I do not think that my thought experiment is too far-fetched.

The point is that contingent value ranking may 1) justify the claim that compassion and fairness are higher than loyalty and sanctity, but 2) still allows, in principle, the justification of possible cruelty and unfairness toward some for the sake of the realization of compassion and fairness toward others and 3) further allows that, if the weighing of less partial values seems even, concerns with more partial values like loyalty and sanctity may also serve as justifications.

This is a way of splitting the difference between a personal and an impersonal ethics. It takes a long view by framing evaluations as provisional and by focusing on

282

gradual change in habits, akin to a virtue ethical focus on character rather than specific actions as the object of evaluation. It establishes the priority of less partial values over more partial values while allowing more partial values to influence ethical deliberation, all else being equal. It thus incorporates one's practical identity, as represented through the expression of more partial values like loyalty, into ethical deliberation without necessarily elevating partiality above impartiality and thereby failing to justify either self-transcendence or communal transcendence. Like the vision of the "perfected heart" from Chapter 5, the ideal deliberative procedure of an ethics for the depressed is underdetermined. I have not gone into detail about how one best proceeds with the weighing of values. But this is a non-ideal theory that does not begin with an ideal model or principle. It instead points toward what may fill in the gaps as one goes. For a depressed person, this is first and foremost the experience of the call of the other.

If the experience of the call of the other is composed of sympathy and alienation, the promise of the call is the opportunity to express sympathy and alleviate alienation. If it seems that sympathy can be expressed in a direct and immediate way, such as to a person suffering in view, a depressed person can act directly and immediately on an intuition rather than having to deliberate and risk spiraling into rumination through hypermoralization. If it seems that alienation, felt as loneliness or isolation, can be alleviated in a direct and immediate way, such as by promptly receiving gratitude or even just being present to see the effect of one's actions on others, a depressed person may be able to act despite demoralization. All this is to say that depressed persons tend toward personal and direct acts of sympathy due to their existential guilt.

But when altruism is impersonal and anonymous, when the recipient is a stranger and the effects are forever unknown, there are more opportunities for depressed persons to lose motivation, not out of a lack of relevant desires or beliefs but out of a doubt or despair that undercuts desires or beliefs at the pre-intentional level. Such failures can also intensify a depressed person's feelings of guilt because they are actual failures to act ethically. The challenge for a depressed person is maintaining motivation in increasingly impersonal and anonymous contexts of ethical action, like giving blood. And the challenge for an ethics for the depressed, which begins by taking the motives as depressed persons as legitimate reasons, is that if a depressed person cannot maintain motivation in impersonal and anonymous contexts, the ethics cannot justify impersonal and anonymous goods. Such an ethics would likely lapse into affirming the priors of whoever practices it. It would then not be recognizable as an ethics.

When I think of my experiences of the call of the other, I think of how it has been as much an obstacle to my goals as it has been a motive for me to act. As a graduate student and instructor, I have often weighed my teaching obligations against my research obligations and found the latter to seem weightless. As is probably obvious by now, I am passionate about philosophical research. But while I would constantly experience doubt in the value of my research and guilt at my pursuing it, I would never doubt the value of helping a student who seemed in need. The more alienated I would become from research, the more I would be inclined to take a kind of solace in helping students: that, at least, seemed valuable. This arrangement, in which my primary obligation was to my research even as the only obligation that felt real to me was to my students, came to seem exploitative. In an era defined by contingent labor, academia thrives on the alienation of

its laborers, who begin by valuing their students and go on doing so through the sacrifice, intentional and not, of much else they value.

I take my impulse to aid my students over doing my research as an exemplar of the call of the other because 1) it seems to speak directly and immediately to me, as an individual, and 2) it arises at once from my sympathy toward the one suffering and my alienation from others and myself, that is, the purposes or goals central to my practical identity. It may really be true that, ethically, helping my students is more important than doing my research, but that is not what the call of the other reveals. The call of the other only reveals that I experience sympathy and that I feel alienated. However, in revealing sympathy and alienation at once, it also may seem to reveal a kind of sympathy that overcomes alienation in motivating action. Thus, although it does not necessarily reveal any ethical truths in and of itself, the call of the other is likely to seem to the one who experiences it like a "ground floor" of ethics: it appears when all that once seemed valuable no longer does. This parallels the sense in which, as discussed in the review of Crowell in Chapter 5, the "call to conscience" arises for Heidegger in the wake of an anxious breakdown.

If "personal and direct altruism" represents the first stage of ethical action for the depressed, the second stage is "personal and indirect altruism," in which the depressed person is aware of or present for the consequences of their actions but anonymous and unlikely to be addressed as the agent. I cherish experiences in which I believe that I have been instrumental in helping others but am able to remain on the margins, so that I am in control of if or how I am engaged by others. I once organized an afterhours dance party with my good friend as the DJ and the face of the operation, so that I could dance in

285

peace while few if any knew that I was responsible for their good times. Over the last year, I have organized a mutual aid project for trans artists in need, with another friend as the selector of the recipients and distributor of the funds so that I may remain anonymous. I have been introduced to the work of interesting artists and musicians, whom I otherwise would likely never have learned of, through this process. I do not experience these acts as involving the call of the other in the sense in which I have defined it, because while some dissatisfaction with the status quo is among my motives, I do not feel my action to stem primarily from alienation. If it did, it would be unlikely that I could sustain the long-term planning required. In alienation, I might succumb to compulsive doubt about the value of these acts or to guilt about their limits. If I were to spiral into allowing the perfect to be the enemy of the good, I would not be capable of the deliberation involved in indirect altruism. Thus, to successfully complete such an act is to take a step beyond the limits set by depression.

Even so, this act, as personal, is experienced as being part of the agent's practical identity, thereby strengthening its motivational power. Specifically, it has the quality of seeming to individuate me, to speak to and of me personally, like the call of the other does. The acts or plans most vulnerable to demoralization or hypermoralization are those that are impersonal.

The third and final stage of ethics for the depressed, then, is represented by "impersonal and indirect altruism," in which both the depressed person and the beneficiary of the depressed person's acts are anonymous to one another. My example of giving blood fits into this category. I have personalized the impersonal to the extent that I can by taking all my motives as reasons in my decision to give blood, but in the end, I

286

cannot count on any sense of belonging or stability emerging from my act, insofar as I have no contact with or awareness of those who may benefit from it. Yet although no sense of stability may emerge from the act, in being able to act in toward impersonal goods, I have become a more stable person.

I mean that in two related senses. First, I have become a more effective agent, more capable of sustained deliberation and action without positive feedback to bolster my resilience. Second, in experiencing myself as a more stable person, I experience the world as a more stable place and its threats as less threatening. My giving blood does not, in the moment, involve an experience of stability (far from it). But the confidence that arises from, and that I may reasonably justify by reference to, my successful act plays a part in increasing my resilience against experiences of precarity overall. This is the concept of character development from the perspective of value ethics rather than virtue ethics.

Recall that an ethics for the depressed begins by taking as normative what a depressed person experiences compulsively: existential guilt, manifesting as sympathy and alienation in the call of the other. Alienation produces a sense of precarity to which a depressed person is averse. Sympathy seems to offer a way of connecting with others that might alleviate that alienation. But the felt tension between alienation and sympathy also represents a practical tension between self-disregard and regard for others. Disregard for oneself and regard for others are both manifestations of existential guilt, but they are in tension because, insofar as self-disregard gives rise to demoralization and hypermoralization, a depressed person becomes practically limited to personal and direct altruism. Because existential guilt is compulsive, this tension cannot be resolved through argument, only through indirect means such as changes in habit. But if a depressed

287

person is motivated by sympathy and the sense of connectedness and stability that may be experienced through it, then the limitations on altruism that are also imposed by existential guilt will seem at best frustrating and at worst, seen through the prism of the guilt itself, contemptible, evidence of one's insincerity or hypocrisy or fundamental unworthiness.

Seen thusly, attaining the capacity to act for the sake of impersonal and indirect altruism is the ultimate experience of liberation for the depressed person from the limits of existential guilt. And, in parallel, attaining the capacity to justify acting for the sake of impersonal and indirect altruism is the ultimate success of an ethics for the depressed. As in my example, although I take all my motives as reasons in deliberation, the motives of impersonal and indirect altruism are not identical with their justifications. I am motivated to give blood in part by pride and the desire for power, at least over myself. But, my own deliberations aside, my reason for action in the sense of the justification I would give in its favor to an interlocutor, is, above all, compassion. I may thus incorporate my motives as reasons, even though those same motives might in other contexts make for bad reasons, so that I may be freed to act and be guided toward valuing and acting on what I judge to be both better motives and better reasons.

This is a pluralism of the heart, in the Schelerian sense of the set of value preferences, that legitimizes the lower, or more partial, values as relevant to ethical deliberation without losing the sense or justification of them as lower than the higher, or less partial, values. As an ethics, it aspires to be reliable and to remain recognizable in the absence of an absolute measure of ethical merit or blame. As a politics, it is melioristic. Though it lacks an ideal deliberative procedure, its deliberations are structured to balance

288

a concern with compassion and fairness above all with a circumspection regarding the possibility of predicting the future consequences of acts intended to realize a more compassionate and fairer world. But, in the absence of a more fully articulated politics for the depressed, I emphasize just one claim that I think an ethics for the depressed would share with a politics: from the value of pluralism follows a resistance to puritanism, defined as the elevation of sanctity above all other values.

In a context of alienation, in which it seems impossible to act in a way that makes a difference that matters, it may be tempting to strive toward spiritual self-purification, to attempt to ensure that all one's motives and desires properly align with one's highest values and principles. This is the strictness of the depressed. It tends to spiral into self-contempt until, either through seeing oneself reflected in others or through hoping to be distracted from oneself by others, it becomes contempt for others. I cited in Chapter 5 the empirical account of disgust as an unreliable moral emotion, one deceptively experienced as certain and absolute when it is in fact adaptive relative to one's upbringing, and I noted the association of disgust with the value of sanctity. This is one reason why an ethics for the depressed assigns sanctity a low value rank. But another reason is to check a depressed person's possible turn toward sanctity instead of sympathy, a turn that I claim is much more likely to fail to sustain a sense of stability.[56]

---

[56] I noted in Chapter 5 that this empirical claim is also, as Haidt's research shows, the default modern American progressive value set: compassion and fairness have intrinsic value, loyalty may have some intrinsic value, and deference to authority and sanctity have no intrinsic value at all. If my position is merely a projection of my political values, that is not relevant to my argument. What is relevant is that, given its resemblance to a standard progressive position, the likely counterarguments are standard conservative positions. I think that a fuller politics for the depressed would be obliged to consider many of these, but here, I consider just one: that communities depend on their members intuitively valuing deference to authority and sanctity for the sake of social cohesion. My response relies on the distinction between the descriptive and the normative that I have sought to preserve: even if that has been the case, it does not follow that it should continue to be the case. Communities premised on disgust of out-groups, for example, might once have truly been necessary for community as such to survive. But this need not imply

A sense of stability, as elation coupled with a sense of connectedness, is basically a sense of belonging. One may experience a sense of belonging while acting in extremely cruel and unfair ways. One may even feel the height of a sense of belonging in such ways, as one affirms one's belonging to an in-group with every wrong committed toward an out-group. In accord with Wong's discussion of flourishing from Chapter 4, I am prepared to affirm that such a person can be both rational and happy so long as they live in a walled garden, shielded from the recognition of their equality with those they wrong. What I do assert, however, is that a depressed person is less capable of living in a walled garden, not because they are intrinsically less cruel but because they cannot count on a consistent sense of belonging. At any moment, for no reason at all, that sense of belonging may seem to disappear. A breakdown may occur, and with it, a depressed person may be thrust into an unsought transcendence. There, seemingly left with nothing, a depressed person will seek something, anything, that still has the power to move them to go on. This is the context in which the call of the other may appear as a kind of rupture or rescue, and thus does possibility of growth, or Schelerian love, manifest. This call is not guaranteed to be as pure as it might seem. It may not actually manifest as a universal love unconstrained by one's group identity or intuitions about who or what is a legitimate object of compassion. It thus would not itself necessarily tend a depressed person toward compassion or higher values generally. What may tend them toward compassion,

that communities premised on disgust remain necessary or desirable today. If this is amounts to a utopian vision of rational self-rule, it is a utopianism by way of meliorism. In its attention to the descriptive and exhortative context of the normative, it conditions a rational community on a trusting, or loving, community.

however, is the eventually habituated anticipation of the call, distinct from the call of the other or the breakdown that may precede it.

I claim that 1) a depressed person is especially inclined toward a breakdown that tends to precede a paradigmatic case of the call of the other, 2) the possibility of breakdown and the call of the other becomes incorporated as part of one's practical identity through conscious awareness and unconscious changes in habit, 3) as the call of the other comes to appear as a personal motive, so sympathy and alienation come to appear as generalizable reasons for action, and 4) this personal transformation amounts to a rehabituation of felt values such that sympathy will seem intuitively to be a higher value than, say, sanctity. I thus attribute a depressed person's inclination toward higher values not to the call of the other itself, but to the dispositions likely to arise from repeated experience of the call of the other and reliance upon it for motivation. This justifies two conclusions. First, the call of the other is experienced through and structured by one's preexisting value paradigms and one's practical identity, and thus it is not "responsiveness as such." However, second, the call of the other may still, at the second-order level of the rehabituation of felt values that it induces, direct a depressed person toward higher values.

But breakdowns are not unique to already depressed persons. To say that any person may suffer a breakdown is to say that not just one's capacity to act, but one's capacity to choose among goals or purposes as a reason for acting, may be lost. One's autonomous will may be undermined by trauma. This is vulnerability, understood as a condition of autonomy.

I thus conclude this chapter with general reflections on the themes of vulnerability and of trust as conditions of autonomy. The extent that an ethics for the depressed will be generalizable to the standard rational agent who is resistant to ethics, like the egoist or the skeptic, will depend on the relevance of vulnerability and trust to ethics, politics, and reasoning in general.

2. Vulnerability, Phenomenology, and Responsiveness as Such

Among normative ethical theories, care ethics calls most attention to the individual's concrete forms of dependence as conditions of the individual's autonomy. Baier (1987) frames care ethics as challenging the "priority" of justice as, in the words of John Rawls, "the 'first' virtue of social institutions" (48). She cites the later theory of Carol Gilligan that an infant may become aware of "two evils": the "evil of detachment or isolation from others whose love one needs" and "the evil of relative powerlessness and weakness." Justice, as the provision of rights to liberty and of security for the weak, may be a response to the latter evil. But an individual may be liberated and secure and yet still isolated. This atomization of the liberal subject is an object of critique from both progressives and conservatives. But care ethics begins as an offshoot of progressive feminist ethics because, as Gilligan observes, "current customs of childrearing" may better attune women who are socialized for childrearing to the former evil of detachment (49). The point is not that women tend to value justice less than men but that women, aware of two evils where men are aware of only one, grasp that justice is not a solution to isolation. Further, this state of affairs is generally presumed not to be innate but rather

292

changeable by ethical thinking that attunes men to the evil of isolation and ethical conduct that counteracts isolation.

Baier remarks that where Gilligan used the language of isolation and care, one might also use the "Marxist language of alienation" and solidarity. She continues:

> the main complaint about the Kantian version of a society with its first virtue justice, construed as respect for equal rights to formal goods such as having contracts kept, due process, equal opportunity including opportunity to participate in political activities leading to policy and law-making, to basic liberties of speech, free association and assembly, religious worship, is that none of these goods do much to ensure that the people who have and mutually respect such rights will have any other relationships to one another than the minimal relationship needed to keep such a "civil society" going. They may well be lonely, driven to suicide, apathetic about their work and about participation in political processes, find their lives meaningless and have no wish to leave offspring to face the same meaningless existence. Their rights, and respect for rights, are quite compatible with very great misery, and misery whose causes are not just individual misfortunes and psychic sickness, but social and moral impoverishment. What Gilligan's older male subjects complain of is precisely this sort of alienation from some dimly glimpsed better possibility for human beings, some richer sort of network of relationships. (51)

Gilligan emphasizes that her "mature women subjects, and some men" did not just want "freely chosen" relationships "between equals" but, as Baier puts it, the kind of relationships "that can obtain between a child and her unchosen mother and father, or between a child and her unchosen older and younger siblings, or indeed between most workers and their unchosen fellow workers, or most citizens and their unchosen fellow citizens." There is a form of "unchosen" dependence that is not just thrust upon them but actively desired, and justice not only fails to respond to its loss but, in its centering of free choice, seems necessarily unable to respond.

It is not enough, Baier notes, for a liberal society to allow some individuals to value justice and others to value care, because such a society would unjustly distribute the necessary burdens of care work. "It will not do," she writes, "just to say 'let this

version of morality be an optional extra. Let us agree on the essential minimum, that is on justice and rights, and let whoever wants to go further and cultivate this more demanding ideal of responsibility and care" (52). First, the cultivating the ideal of care requires cooperation, not just noninterference. Second, letting "whoever wants to" perform care work risks creating and intensifying exploitative relationships because these desires may be the products of socialization. It is itself unjust to watch placidly as women, socialized to identify care work as required for the successful performance of their gender roles, shoulder greater burdens and then shrug at the fait accompli. Third, care work or reproductive labor is necessary for the continuation of a society, just or not. "Moral theory," Baier concludes, "cannot regard concern for new and future persons as an optional charity left for those with a taste for it. If the morality the theory endorses is to sustain itself, it must provide for its own continuers" (55). In doing so, moral theory will recognize the distinct importance of unchosen relationships and "contract" will cease "to seem the paradigm source of moral obligation," just as justice will seem at best only equal to care.

For an ethics for the depressed, it is essential to recognize that isolation is a problem distinct from powerlessness or weakness and that care is a solution distinct from liberty or security. In existential guilt, the sense of precarity and of isolation may seem to be identical. It may seem at once that because one is isolated, one's motivation is precarious, and that because one's motivation is precarious, one cannot cooperate with others. But although a depressed person may feel themselves to struggle with one all-encompassing problem, it can be seen to require at least two solutions: justice and care. Similarly, the contingent value ranking of Chapter 5 frames fairness and kindness as

294

equal "less partial" values. I suspect that care ethics will seem intuitively correct to depressed persons, who may be and see themselves as something like "aspiringly autonomous," whereas already autonomous individuals may require persuading. It is thus not coincidental that the move from personal to impersonal goods that I enact in this chapter parallels the move from dependence to autonomy. As I understand them, a depressed person is not a child: they are already reasoning at an adult level. But their status as aspiringly autonomous means that some similar progress to full autonomy is required, one that is more likely to involve therapy or treatment than education. Still, if the theory of an ethics for the depressed is anything, it is the "educational component" of that movement to autonomy, for whatever it is worth.

As I discussed in Chapter 5, in my identification of felt values compatible with depressive reasoning and, contrastingly, the structures of demoralization and hypermoralized deliberation that undermine the autonomous reasoning of depressed persons, I have relied on an unorthodox "critical" form of phenomenology. This phenomenological method resembles, and is inspired by, Beata Stawarska's (2018) discussion of feminist phenomenology. As Stawarska argues that a non-transcendental phenomenological method allows the phenomenologist to attend to the operation of mechanisms of social power operating within phenomenology itself, so I argue that a non-transcendental method is necessary for the depressed person to distinguish, within the "given," between compulsive intuitions that tend toward alienation from one's values and non-compulsive intuitions that tend toward engagement with and through one's values.

In her account of feminist phenomenology, Stawarska shows the consonances between phenomenology and feminist philosophy. Phenomenology's study of "human reality in a concrete sense, as it appears to the experiencing subject, emphasizing the bodily and socially modulated quality of lived experience and its expression" is suited to "the feminist project of making women's historically devalued experiences visible" (13-14). For example, where a historically male-dominated discipline of medicine may take the pregnant bodies of women to be objects of clinical study, feminist phenomenology may disclose and elevate a distinct domain of knowledge about experiences of pregnancy that would otherwise be obscured or suppressed. The collaborative feminist effort to recognize and express this phenomenological domain may be "taken up by consciousness-raising groups" that help "women to actively assume the role of knowing subjects reflecting on and interpreting their condition," which in turn may be "a source of empowerment for women's groups and may lead to women assuming a more agentive relation to their bodies" (14). The recognition and expression of women's experiences in a mutually supportive context may assist in developing their effective agency.

But, she argues, the method of transcendental phenomenology introduces dissonance by claiming to describe invariant transcendental structures that are not contingent on empirical or historical context, which amounts to a claim to an objective standpoint which feminist philosophers may reject or, at least, find dubiously reminiscent of masculine philosophical presumptuousness. Formally, phenomenology has "historically assumed the category of purportedly neutral and ahistorical existence as its field of study and focused on the constitution of meanings within the sphere of consciousness purified of contingent empirical content." Not coincidentally, then,

296

"categories of gender and sexual difference were not counted together with a temporal and intentional structure among the essential features of consciousness" (14). Some phenomenologists, like Levinas, correlate femininity with an ontological or transcendental frivolity, seeming to project masculine bias into what he is ostensibly observing as the invariant structures of consciousness. Stawarska refers to Luce Irigaray's argument that "feminine eroticism is envisioned by Levinas as a threat of excessive materiality to the ethical relation of transcendence enjoyed by a masculine subject." This sounds like a man frustrated that his thoughts of women are distracting him from philosophy, not an impartial phenomenologist.

Stawarska acknowledges that these facts may motivate feminist skepticism of phenomenology. Transcendental phenomenology may attempt to "identify a set of invariant essences within human experience at the risk of presenting a stock of gender stereotypes in the guise of universal and unchanging truths." Not only does this belittle women and femininity, but it precludes what is arguably a central goal of feminist philosophy: "an analysis of the ideological motivations and power structures leading to such stereotyped notions" (15). If the transcendental phenomenologist believes themselves to have objective insight into invariant structures of consciousness, they will likely also believe themselves to be immune to the effects of ideology or socialization, at least when doing phenomenology. The phenomenologist may think that this must be possible insofar as phenomenology is the essential method for the essential task of identifying invariant structures of consciousness. But a skeptical feminist is likely to view their position to be naïve at best and their method to be doomed.

Stawarska defends the possibility of a feminist mode of phenomenology that, by not being fully transcendentally "purified," may be cognizant of the operations of "mechanisms of social power" (18). Grant that a phenomenologist, in turning to reflection on what is given in experience, may reenact ideology or socialization and so, like Levinas, mistake an individual (if socially normative) bias to be a generalizable feature of consciousness. Even so, Stawarska responds, "subjective experience is not simply a by-product of dominant social arrangements": it "provides also a site of critical reflection, resistance, and revolt" (20). Experience is not merely a theater where plays staged by power perform. In attending to the features of experience, one may notice features of power which no concept yet adequately names. For example, the concept of "sexual harassment" is a relatively recent feminist coinage that responds to and accounts for a possible feature of experience, one for which neither "bullying" nor "flirting" are entirely adequate descriptors. Sexual harassment may be insidious in its creation of hostile environments, and it is even more insidious when it is unnamed or, in the absence of the concept, unnamable.

In Kantian terms, both concepts and intuitions have parts to play in the development of philosophy. If linguistic or historical analysis does the work of attending primarily to concepts, and phenomenology does the work of attending primarily to intuitions, then philosophical collaboration seems both possible and advisable.

To detail a feminist method of phenomenology, Stawarska adapts Simone de Beauvoir's phenomenology into a method that oscillates, to borrow a term from her fellow Beauvoirian Bonnie Mann (2018), between the concreteness of social and historical phenomena and the abstraction of transcendental and existential categories. "By

298

approaching gendered identity on its own terms," Stawarska argues, "Beauvoir tackles a phenomenon which depends to a degree on a subjective constitution of meaning (becoming a woman as a freely undertaken project in the existentialist sense) and which is informed by the "total situation" of social structures, power disparities, and dominant gender norms (e.g., the ideological construct of feminine mystery with its concomitant economic dependency and social subordination)." As the quintessential feminist existential phenomenologist, Beauvoir combines an existentialist view of the feminine subject as ultimately free and fully responsible with a phenomenological view of the feminine subject as "*a situated subject* who undergoes as well as constitutes meanings, and whose freedom is enabled and constrained by a social-structural positioning in the world." For example, a woman's subjective experience of herself as mysterious to men may be better understood not as an awareness of some essential feminine mystery structuring her self-consciousness but rather as arising from "the accumulated history of devaluing women's speech in a misogynist society" (24). Feminist phenomenology in the style of Beauvoir does not focus exclusively on the "subjective constitution" of "meaning-making acts" but also on the subject as "the effect of multiple social norms and conventions" (25). To understand the phenomenon of becoming a woman, it is not enough to understand sociality from a transcendental standpoint. A phenomenologist must become able to see their own practice as emerging from their socially and historically situated position and, to mitigate bias, disavow the phenomenological method's claim to transcendental authority in its exchanges with empirical methods.

Much as I develop the concepts of the "qualitatively ambiguous situation" and the "qualitatively opaque problem" to resolve the classical and neo-pragmatic disagreement

299

over the relevance of lived experience to an account of normativity in Chapter 3, so Stawarska's account of non-purified feminist phenomenology aims to resolve the phenomenological and skeptical feminist disagreement over the relevance of lived experience to feminist normative projects. I think that this normative question about lived experience, which shows up in both pragmatism and phenomenology, resembles a depressed person's problem with normativity. The latter is, as I argued in my introduction, the problem that motivation or "feeling justified" presents itself as a greater normative problem than justification or "being justified." A depressed person's feelings of being "unable to go on," in rendering their existing reasons for going on seemingly impotent or inert, may resemble a lack of permission or justification for going on, insofar as the problem may seem to them to be their having inadequate reasons for going on rather than their having a motivational problem. This is fundamentally a compulsive confusion of motivation and justification and, thus, a mistake about how normativity or justification works.

But it reveals something important at and about the margins of philosophy: the extent to which justification seems like a first-order problem may depend on the extent to which one sees oneself as, at least potentially, included in the "space of reasons" where justification takes place. In this light, I would suggest that a depressed person, rather than (or more than) being confused, realizes something: before they can deliberate effectively alongside others, they must solve the motivational problem that leads them to ruminate rather than deliberate effectively. Their view of normativity is unorthodox because they are, or feel themselves to be, excluded from orthodox normative discourse, even when that is discourse with themselves. But where depressed persons may be mistaken about

300

their exclusion because their feelings are the product of compulsion, a feminist critic may correctly assess that they have actively been excluded from the spaces where norms are decided. From that vantage, one might reasonably argue that justification is a second-order problem that follows the now first-order and intrinsically philosophical problem of, in Stawarska's phrase, "social-structural positioning" within and around the space of reasons.

Christinia Landry's (2018) discussion of Beauvoir's ethics of ambiguity locates an advancement on Sartre's or Levinas's understanding of alterity that is absent from Crowell's account of the origins of responsiveness, discussed in Chapter 5. For Beauvoir, Landry argues, direct conflict or a dramatic encounter with the other is not necessary to become an object to oneself. Instead, the other may solidify within one slowly over time as, for example, one's body becomes more an object for others than expressive of oneself. This occurs when "shame is introduced through sexual maturation … In becoming flesh for the Other, the little girl realizes that she is no longer a subject of action; her world, once an open set of possibilities, begins to close down." As the girl becomes a woman who is objectified as and for her body as childbearing and sexually desirable, "her subjectivity (her possibilities for transcendence and thus for freedom) emerges as bound up in the ways in which she *is* for the other" (93). But the gaze of the other is not necessarily objectifying or shame-inducing, as it seems to be in Sartre's *Being and Nothingness*. For Beauvoir, there is a "possibility for collective disclosure," as when one admires the sky or a snowfield and one feels and accepts a sense of distance from what is disclosed in the experience, understanding that "one cannot be wed with that which one contemplates" (89). The possibility of shame-inducing or -inflicting

301

objectification by the gaze of the other, insofar as objectification is a denial of equal subjectivity, is a problem for collective disclosure between subjects. But this problem may be resolved through, as Landry says, an "intersubjective eroticism" or "a promiscuity between gazers" in which "one recognizes that one is *for* the Other as the Other is *for* oneself" (94). An experience of freedom through dominance of the other as one's object is supplanted by an experience of ethical freedom through collaborative disclosure with the other as a fellow subject. Similarly, the experience of sexuality as a source of shame gives way to an experience of sexuality as a source of freedom.

The basis of this affective or erotic sense of collective disclosure with another subject equal in freedom to oneself is vulnerability. In Landry's words:

> We are vulnerable through our gazing insofar as we acknowledge that in gazing we are also gazed upon; our projects, the way in which they are disclosed, and the freedom required to take them up are also under consideration. … This intersubjective eroticism—a messy encroachment of affective beings—is rooted in the body as a body that can both see and be seen, touch and be touched— disclose and be disclosed; it is a body that is an "openness" onto the world. (95)

Phenomenologically, vulnerability is the possibility of becoming an object to oneself. The experience of self-objectification can negate freedom, as when one is restricted from achieving one's goals by the anticipated prospect of shame or punishment, either inflicted by others or, due to the internalization of oppressive norms, inflicted upon oneself. But the experience of self-objectification is also compatible with the experience of one's subjectivity and, similarly, the experience of objectifying another is compatible with the experience of their subjectivity. In Landry's context of eroticism, it is possible to find only joy, and no shame, in freely being an object to another and in another freely being an object to oneself.

Ethically, vulnerability is the concrete possibility of freedom or autonomy and, thus, the possibility of its loss. The preceding discussion of care ethics has emphasized the necessity of relationships of dependence in developing autonomy. But as autonomy is developed through care work or reproductive labor, so autonomy is sustained by norms against exploiting the vulnerability of others. As Baier says, these are norms of care as well as justice. One may feel isolated even among mutually respectful equals because the cure for isolation is not company but dependence: that is, allowing oneself to depend on others and allowing others to depend on oneself. In turn, this dependence is not ideally symmetrical, as in a contract or fair exchange, but necessarily asymmetrical even between equals, arising from a need that care fulfills. The continuation of society may, following Baier, depend on more of its members accepting responsibility for relations of dependence and reciprocating, over time, asymmetrical care relationships, just as parents care for children who then care for aged parents. Similarly, the sustenance or flourishing of the individual may depend on the recognition of vulnerability, in both oneself and others, as a motive of and reason for responsiveness to others.

Crowell's reading of Heidegger's concept of the call to conscience and Levinas's concept of the face as "responsiveness as such" partially dovetails with the account of vulnerability. As transcendental phenomenology, it excludes an account of the social processes of internalization of concrete modes of responsiveness as may be found in Beauvoirian phenomenology. This exclusion, exposed to a feminist perspective, in turn reveals an incompatibility of content, not just of form, between Crowell's account and an ethics for the depressed.

The formal problem is, as I suggested in Chapter 5, that judgment as employed in the transcendental phenomenological method is generally insufficient for a depressed person's distinguishing between other-regarding and self-disregarding intuitions. Indirect methods like treatment or mental techniques are typically necessary to defuse the downward spiral that may follow from intuitive self-disregard. Transcendental phenomenology depends on a capacity to judge the difference between invariant and variant structures of consciousness, a capacity that a depressed person will likely doubt to be reliable. Moreover, they will likely take the presumption of its reliability to be dangerous because spiral-enabling. This concern with the compulsive bias of judgment parallels a feminist concern with the mechanisms of social power compelling biased judgment: both reveal a form of dependence or vulnerability that is obscured by the phenomenological presumption of absolute autonomy or authority.[57]

The problem of content is that the other revealed by the call is vague and, thus, an ethics founded on responsiveness to the call of the other will be excessively vague. I affirm Crowell's reading that the face, for example, represents a coming to awareness of the possibility of responsibility as such. Thus, it is understandable that the face as a concept does not have concrete content, or in other words, that it is no one's face in particular: if it had this content, it might just evoke responsibility for a particular person with that face, not responsibility as such. This is just to say that the face is vague on purpose. Its vagueness arises from its unique phenomenological position as, like the call

---

[57] Ironically, this critique parallels Scheler's critique of the *epoché*, also discussed in Chapter 5: because it depends on judgment, and the preferring of values precedes judgment, the reduction cannot reveal the subjective ordering of values. The irony of this parallel is that Scheler thought that a phenomenological attitude, as a kind of "spiritual seeing," could observe an objective order of values. That is, it would be an *even purer* method, not less so.

to conscience, emerging from a space of normative breakdown. Yet in constructing a normative ethics, we soon run face-first, as it were, into a basic question: "How do we know who has a face?" If the answer depends on some intuition, then given the previous challenge to the transcendental phenomenological method, the supposed necessity of ethical responsiveness in concrete situations immediately begins to teeter on a questionable foundation. Phenomenological intuitions no longer seem quite so trustworthy. But the answer presumably cannot depend instead on some formal criterion of face-ness, as the successful application of this criterion would involve the content of the concept of the face: something like, "Well, the face has X characteristics, and Y has those characteristics, so Y must be a face."[58] So be it, the phenomenologist may reply: ethics is vague.

But ethics is not that vague. There is more in the way of content that an ethics for the depressed can offer to the depressed than a vague gesture at a transcendental condition of normativity. I assert that the call of the other, the same one that Heidegger calls the call to conscience and Levinas calls the face, *just is* sympathy, alienation, and the experience of sympathy as an individuating solution to alienation. This ethical phenomenon is important because it arises in contexts of normative breakdown, but it remains psychologically simple. This simplicity is obscured by the theoretical attempt to

---

[58] As in the case of Scheler's "essential criteria" for values, a turn toward a formal criterion for alterity would also be a turn toward transcendental philosophy in the style of Kant and away from transcendental phenomenology. It is not obvious why the experience of an encounter would be necessary if such a criterion could exist. That said, I do not claim here to have refuted Levinas's ethics, as I have not represented it in sufficient detail to do so. I see myself just as describing an essential problem to which I provide a different answer that I think may justify more in the way of ethical content. I am aware that some commentators, reacting to his writing on "Zionism and the Palestinians," have challenged Levinas's philosophy of alterity and his arguably underdetermined concept of the face as "too unsafe as a prescription for a political order" (Caro 2009, 671). But others have defended Levinas as repeatedly misinterpreted on this point by commentators (Eisenstadt and Katz 2016). I am not prepared to say who is correct.

position it as "responsiveness as such." This is, to me, a phenomenological failure as well as an ethical one.

I am sympathetic to the elevation of the call of the other as being the experience of finding an absolute source or foundation for ethical motivation and reasons. It would be nice to have an absolute source or foundation for ethical motivation and reasons. But if I am right that the call of the other definitionally involves an experience of alienation, then to claim that the call of the other *really is* the foundation of ethical responsiveness seems to me, in the context of depression, to restrict the scope of ethical action to responses to alienation. This is, at best, a tragic and frustratingly limited view of ethical possibility and, at worst, politically reactionary if alienation is a response to social progress.

More to the point, if the call of the other is a concrete phenomenon, then the claim is just false. My placing my students' needs above my research obligations is a response to a specific experience of the call of the other, but my giving blood is not. In giving blood, I neither feel myself to act from, or in response to, alienation, nor do I have any sense of the presence of the others whom I seek to aid: these strangers are totally abstract to me. If one counters that the call of the other is not an episodic experience or encounter, if one insists instead that it is an invariant structure of consciousness that *must* be present even in my giving blood, then I just do not see the connection between this ontological abstraction or spooky metaphor and the concrete experience that I am describing as the call of the other. Ontology becomes unmoored from phenomenology in theoretical ploys like this. Not every condition of normativity, necessary or otherwise, is an invariant transcendental structure, of consciousness or otherwise.

The comparatively simple answer is that the call of the other is not a necessary condition of normative responsiveness: it is a contingent condition of normative responsiveness that is especially relevant in the context of anxious breakdowns and that may be experienced, in that context, as though it were a necessary condition of normative responsiveness. Phenomenology is supposed to be able to distinguish necessary from contingent structures of consciousness. Perhaps if phenomenology at large followed the lead of feminist and critical phenomenology and acknowledged its limitations, it would successfully do so more consistently. I say this not to provoke any interlocutors but to emphasize the possibilities of phenomenology in identifying the structures of ethically significant phenomena like demoralization, hypermoralization, and the call of the other. Without phenomenology, experiences of depression remain obscure. Depressed persons thus cannot afford for phenomenology itself to be obscure or obscurantist. I have argued that an ethics for the depressed is a partial ethics but that it must justify the priority of impartial values. Similarly, an ethics for the depressed begins with experiences of ineliminable vagueness or ambiguity, like existential guilt, but it justifies its value in the clarity it brings.

The call of the other is one contingent way of revealing vulnerability: in an experience of extreme isolation like an anxious breakdown, an act of care for another may appear as a solution. According to Baier's account of care ethics, the call of the other reveals a specific truth, even if, as I claim, it would not be ideal to restrict ethical action to responses to the call of the other. That being: accepting the dependence of a vulnerable other or having one's own dependence be accepted is, more so than respect from an equal, an antidote to loneliness. As the feminist account shows, vulnerability may be

revealed otherwise, as through the shameful consciousness of one's embodiment during and after sexual maturation, or through the joyful consciousness of one's being cared for in consciousness raising alongside friends and allies. I turn now to the latter positive response to vulnerability, which I broadly categorize as trust, and its ethical and political possibilities. For reasons of space, I will keep my discussion of trust relatively brief.

3. Trust, Self-Trust, and Making a Trustworthy World

On trust, Baier (1986) observes: "We inhabit a climate of trust as we inhabit an atmosphere and notice it as we notice air, only when it becomes scarce or polluted" (234). I have come to take the phrasing of "atmosphere" as a clue to the presence of a pre-intentional state, one seemingly directed at everything or nothing in particular. Her first provisional definition of trust is "accepted vulnerability to another's possible but not expected ill will (or lack of good will) toward one" (235). As pre-intentional, this acceptance may not be represented by a first-order belief, at least initially. "The natural order of consciousness and self-consciousness of trust," Baier says, "progresses from initially unself-conscious trust to awareness of risk along with confidence that it is a good risk, on to some realization of why we are taking this particular risk, and eventually to some evaluation of what we may generally gain and what we may lose from the willingness to take such risk." As in the care ethical understanding of dependence, the initial form of trust is infantile, easing and obscuring Gilligan's evil of isolation. But "the ultimate point of what we are doing when we trust may be the last thing we come to realize" (236). Willing and reasoned trust may be the foundation of a just and caring social world. Still, it is possible to exploit even willing and reasoned trust. There must,

Baier concludes, be a test of moral decency of trust. The challenge is that, like a medical examination that requires an invasive procedure, to deliberately test the decency of trust may mean to damage it. Even so, it remains at least theoretically possible to make a more trustworthy world, where willing and reasoned trust is more reasonable to will and less likely to be exploited. I anticipate that a politics for the depressed would proceed along such lines.

What Baier calls a "climate" or "atmosphere" of trust is what Jay Bernstein (2015), following Baier, calls "trust in the world" (110). Bernstein reads Jean Améry's phenomenology of torture as expressing an example of "devastation," referring to "the experience of being undone in one's standing as a human, and to continue to experience the event of destruction as a present moral fact about oneself" (75-76). This recurrence of harm in devastation is an element of trauma, which, in making present in one's own experience what were past harms, induces a general sense of precarity and isolation. Bernstein takes this possibility of being traumatized to reflect a general condition of "existential helplessness," epitomized by the "involuntary body" (122). Those who are accustomed to experiencing their own body as a vessel or a medium for their willing action may be struck by shame at an experience of their body's involuntary processes taking over, as in, say, incontinence. This shame is not just feared or anticipated from others but arises from a sense of shock and betrayal at the possibility of one's body to resist or exceed one's will, as in my feelings over my phobia associated with giving blood. Part of what it means to experience oneself as an agent, or as willing for oneself, is

to reconcile, in Bernstein's terms, being a body with having a body in one's self-understanding.[59]

Bernstein frames Baier's definition of trust as a necessary and universal requirement due to the necessary universality of vulnerability for human beings. As with Baier's concern regarding the prospect of a test for the decency of trust, Bernstein notes that trust must operate unobtrusively. It cannot be generated at will and so cannot necessarily be restored if destroyed. The generation of trust, in Bernstein's developmental account, begins with the first love of parents. The infant who fears powerlessness and isolation becomes an adult who trusts implicitly, without asserting or even thinking it explicitly, that they are neither helpless nor alone. But this trust is only trust, not certainty. Should this adult find themselves helpless and alone, as did Améry as the victim of torture, this self-trust may be shattered forever. They will relate to themselves fundamentally differently, paranoid about how their own minds and bodies may turn or be turned against them.

Self-trust is trust in one's own dignity, which Bernstein summarizes as a cluster concept or "the dignity constellation" (265) of respect, self-respect, love, and lovability. After first love comes self-respect, the expectation of appropriate treatment by others, which is also to believe that one is lovable by others. Bernstein's observation, inspired by Améry, that there is an embodied aspect of dignity, involving the appropriate relation

---

[59] I elide for now Bernstein's Hegelian-Brandomian account of self-consciousness. In his review, Craig Duncan (2016) presents as overly reductive Bernstein's claim "according to which the self is a normative construction constituted through its relations to others, so that recognition of the right kind by others is necessary for one even to exist as a self." I would discuss this account of self-consciousness further in a fuller discussion of practical identity. But I suspect that Bernstein's argument succumbs to a problem that parallels that of Crowell's: it may involve the projection of a contingent possibility or variant transcendental structure as a necessary and invariant structure.

between the body one has and the body one is, motivates his claim that the moral horror of the treatment of the bodies of Jewish victims of concentration camps arises from a sense that devastation, an attack on dignity, has been inflicted upon them. This in turn, as in phenomenological value ethics, motivates a critique of Kantian or legalistic ethics: wrong lies first in wronging persons, not breaking rules.

This personalistic politics resonates with Jill Stauffer's account of ethical loneliness. She too is inspired by Améry, arguing that he experiences a doubled devastation: he first experiences himself as helpless and alone in torture and then, in experiencing his sufferings as unrecognized and denied by others, experiences himself as helpless and alone again. "The experience of having been abandoned by humanity," Stauffer says, is "compounded by the experience of not being heard" (2015, 1). To counteract ethical loneliness requires understanding the doubled nature of its devastation. Stauffer recommends cultivating a practice of listening as much to the unsaid as what is said, along with increased awareness "of the fragility of human safety" (110) or vulnerability. But she goes further: not only is justice insufficient to counteract ethical loneliness, as Baier would affirm, but for the recognition of vulnerability by complacent autonomous subjects, "some destruction of their idea of the self's autonomy" (81) might be necessary. In the terms of an ethics for the depressed, people may need to be alienated from their walled gardens to be motivated to pursue transcendence and form solidarity aligned with higher values.

For depressed persons, I predict that the call of the other will be enough. The call of the other is a contingent but, in the context of normative breakdown, relatively reliable motive for action that, in seeming to be an individuating response to alienation, is easily

311

taken to be a reason for action. I have argued that depressed persons, in being especially susceptible to normative breakdown through the loss of a sense of stability or belonging and the emergence of a sense of precarity or isolation, may and do rely on the call of the other to motivate personal and direct altruism. But already autonomous persons who may rely on a more consistent sense of stability or belonging will be less likely to experience the call of the other as I have defined it. Even if they experience sympathy, which I presume most persons do, that sympathy may seem unproblematically to them to apply only to their in-group. Alienation is not a given for persons in general the way it is for the depressed as I understand them. What I have called the walled garden is, essentially, the negation of the possibility of alienation. Therefore, I conclude that while the call of the other supplies the necessity motive and reason for self-transcendence that completes an ethics for the depressed and makes it a recognizable ethics, it is insufficient for generalizing an ethics for the depressed into a value ethics of engagement for already rational persons, including egoists, skeptics, and those who dwell in walled gardens. The intuitive value rankings of an egoist, for example, might not contain within them any reason for self-transcendence.

The solution, I have proposed, is an eventual politics for the depressed. The limit of ethics, I think, is the building of trust: trust in oneself and trust in others, despite and because of our vulnerability. But ethics alone is insufficient to build a trustworthy world. Trust in an untrustworthy world is probably unreasonable and, more importantly, surely unfair. It is unfair to require vulnerable persons to trust in communities or institutions that relentlessly exploit them. But society cannot go on without trust or care. The answer, then, is not to require vulnerable persons to tacitly accept their exploitation, but to

motivate and justify, from within ethics itself, the endeavor to eliminate exploitation and to realign society in terms of higher values.

Of course, to justify doing so is not to prove that it is empirically possible or to compel anyone to do so. But, as any impartial ethicist would say, to reasonable persons, justification matters. It matters to be able to legitimately claim that actions that reject lower values in accordance with higher values are right. It matters to be able to legitimately claim that actions that reject higher values in accordance with lower values are wrong. This is the hope of ethical theory itself, framed in terms of value ethics. I imagine a politics for the depressed as not just a broader defense of the depressed, but a broader defense of this existential hope for ethics.

I conclude my dissertation with a prologue to a politics for the depressed, one that does not require trust in one's adversaries but that nevertheless justifies hope that they will change.

As Stauffer says, people may need to be alienated to form solidarity aligned with higher values. Her term "destruction" might suggest that this is inflicted through, say, confrontation via protest or the exertion of leverage via withholding labor or striking. This seems plausible. There are, as the example of Améry shows, also more violent alternatives for undermining self-trust or illusions of autonomy. But there may be milder or more hopeful ways of inducing a reevaluation of one's own autonomy or sense thereof, too. In my final section, on what I call "proto-politics," I consider one: an act motivated and justified by integrity which, in risking alienation from a community, unintentionally leads to new solidarity.

4. Proto-Politics as Integrity and Solidarity, or: My Family

I hesitated to write this final section. First, the dissertation is long enough. Second, it functions as a prologue to a politics for the depressed, which I cannot adequately develop without more research into practical identity and political philosophy. Third, the example strikes me as potentially disloyal. I was willing to write about my mother in the introduction because she passed away almost twenty years ago, and even references that risk diminishing her are likely still to honor her. My father is alive today. Even if he would not be hurt by what I have to say, I am instinctively reluctant to do anything resembling airing our family's dirty laundry.

But today, I learned that my stepmother has been in a car accident and, as I write this, I do not know her condition.[60] So, I have decided to take the risk of acting in a manner that may resemble betrayal to express what, to my mind, honors my family. This is the story of the most important argument that I ever had with my father and, to me, the most important conversation that I ever had with my stepmother. It is not yet "politics," but it is, I think, "proto-politics."

For as long as I can remember, my father has had difficulty intuiting the social expectations of others. He is an intelligent man with a logical mind and an expressive

---

[60] I am adding this footnote later to ease my reader's possible concerns: she survived without cognitive or serious spinal injury but requires ankle surgery and must wear a back brace for a lumbar fracture. She was the victim of another driver who was speeding and alone in his car. She also sustained several other bone fractures due to the seatbelt and driver's side air bag that, as my father observed, almost certainly saved her life. Tragically, the other driver perished at the scene. I do not know his identity. Though I have no way of reaching them directly, I offer his family my sincere condolences. To my fellow citizens: let us work toward changing our cities to have more public transportation and fewer cars. Doing so earlier may not have prevented this crash, but it might preclude the next.

heart. His reasoning is often sound, and his sentiments are often beautiful. For years, his primary hobbies have been gardening and writing poetic couplets, often riffing off idiomatic phrases, and sharing them via email. But his assumptions about how others will react to him have often been mistaken, and his thoughts about what they will or should find appropriate have thus gone astray.

One day, at my cousin's wedding, my father had a brief and curt interaction with the groom. One might expect that my father would contextualize this interaction with the understanding that the groom had met many new people from his bride's family that day and was likely distracted or fatigued. I think that, unfortunately, my father did not. Instead, he chose to, as is his wont, write an email about his experience of the encounter and share it with his usual audience, which included our family members and some of his friends who, I believe, are strangers to the family. To my father, this was likely intended primarily as personal reflection, though perhaps a part of him also wanted to vent about an unexpectedly negative interaction.

If you, the reader, have typical social expectations or experience typical social intuitions, you may be groaning already. Yes, my uncle, the bride's father, was, as I recall, enraged. My father had told a belittling tale about the groom, newly a member of our family, in a message addressed to strangers that included family members. A reasonable interpretation of this act would be that my father intended to put a family member, whose position in the family might not yet be settled, on blast while passive-aggressively notifying the family of his deliberate act of malicious betrayal. After all, why else would someone write an email like this? What sort of person would, in good

faith, find this apparent complaint to be a basically innocuous act of self-expression and, furthermore, care enough about self-expression to even risk possible offense?

The answer, of course, would be my father. But his subsequent behavior would only strengthen a case, no doubt representing my uncle's position, that he was acting in bad faith.

I think that I was around twenty-four when this happened, which would mean that my sister was in high school, and I happened to visit around this time. If memory serves, I was informed by my stepmother that we would be taking an unexpected trip to visit my paternal grandmother in a nearby city. I thought nothing of it, as we did this often when I visited. I also thought nothing of the folder of documents that my father brought with him, as he is a very meticulous person who often prints up itineraries for even brief ventures. Only when we arrived did I realize our true purpose. My father, realizing that the family had turned against him, had decided that he would take the initiative to visit his mother and tell his side of the story, complete with printouts of the relevant emails.

I remember my father beginning to give his presentation. I remember my sister beginning to cry in what I interpreted as humiliation before I stormed out of the door of my grandmother's townhouse. I don't remember my stepmother's expression. I was so angry. As a teenager, I had frequently been embarrassed by my father's behavior and, even more acutely, frustrated by my difficulty explaining to him why I was embarrassed. That day, in my eyes, my father had started losing a fight and then ran to mother. It must be nice to be able to run to mother.

I decided that I would finally win an argument with my father. This decision came despite my judgment that it was practically impossible to win this argument. Being very

316

logical, he would target any weaknesses in my reasoning. More importantly, he simply

did not share my typical social expectations or intuitions. To win the argument, I would

have to not only perform ironclad argumentation while very distressed but also convince

him that my social intuitions were accurate. I could try to do that by revealing

discrepancies between his social expectations and the results or by revealing internal

tensions in his own thought process. But ultimately, he would have to trust me, even as I

confronted him in anger. I concluded that I had no hope of succeeding. But it didn't

matter to me. I felt that my essence rebelled against him. I had to speak.

In the car ride back, with my father driving and my stepmother in the passenger

seat, I informed him that I objected to what he had done and that I wanted to refute the

reasoning behind his actions. I also informed him – I recall him seeming a little

bewildered – that I would likely lose control of my emotions at times during the

argument and that, during that time, I would ask for a break for me to compose myself

before resuming. He agreed to these terms.

We had to take a few of those breaks. But my strategies for success were

effective. I explained how the family likely saw his behavior and challenged, among

other points, that he had chosen to send the email "to" his friend but "CC" the family.

When he objected that the distinction meant nothing, I asked why he had then made the

distinction, to which I recall him struggling to respond. I think that I pushed him to a

breaking point, because he concluded with a striking claim that I had never heard him say

outright but that seemed consistent with the way he lived. He said that he believed

himself to be responsible for his choices but not the social consequences of his actions,

because the reactions of others were fundamentally unpredictable and, thus, it would be

317

unfair to hold him responsible. I countered with what may have been an elitist point but one I think was nevertheless effective: even the most vulgar and unsophisticated gossip acts based on predictions about the social consequences of their actions, like the harm that they expect to cause. Most people, I asserted, can predict the responses of others with better or worse success. To these people, his claim to be responsible for his choices but not others' reactions to those choices will just seem like a total disavowal of social responsibility.

My sister, wisely, had put on headphones. I don't recall my stepmother saying a word.

When we got home, I, exhausted, went back to my room. I think that I felt empty yet satisfied in that emptiness. I had achieved nothing, I thought. My actions would have no consequences. Yet I had, in a blaze of passion and rhetorical ability, represented something like my essence to my father, a person who had formed me. I felt fully myself.

Unexpectedly, my stepmother soon came to visit. My stepmother is an immigrant and, as a teenager, I had difficulty interpreting the meaning of her statements. I recall her once saying to me with a kind of desperation, "It's alright with me if you clean your room." I was confused by the meaning of this for longer than I care to admit and then, I think, resentful of what I saw as her passive-aggressiveness. I now respect what must have been the effort involved in her being as direct with me as she could be, contravening her view of indirectness as basic to courtesy.

By this point in my life, I had learned how to better listen to her. I recall being mostly quiet as she talked about various things. But when she got to her point, it stunned me. She said that she had always assumed that my father, as an American, understood

Americans better than she did. Even when she thought that his explanations of the behaviors of others couldn't possibly be correct, she concluded that he must know better. But now she thought that, all along, she must have known better. Even a gossip, she said, better understands these things than him.

I had been so consumed by my climactic confrontation with my father that I had forgotten she was in the car. But she had been listening. She had been thinking the whole time.

I had experienced my act as arising purely from integrity. I had felt, as when I confronted my mother days before her suicide to take her wine from her and pour it into the sewer, that if I did not act, I would simply disappear, or somehow cease to be myself. But I had failed to predict, due to my immaturity and limited knowledge, what the meaning of my act would end up being. I am ashamed now of how foolish and arrogant, even sexist and racist, it was of me to forget her presence during my honorable duel with Dad. But I value what I think is the most important lesson of this story, besides that I should better attend to the presence of women during my confrontations with other men. That being: I had first needed to trust that my integrity was worth acting on for my act to have the unintended effect that was its true significance.

It may seem strange to call this story one of "proto-politics" for anyone, let alone the depressed. A family argument is hardly what I imagine when I think of politics, especially one so rigorously and fairly conducted. Unfortunately, I currently have little experience with political activism or participation beyond voting and reading to prepare for voting, so I have few political examples to offer. At the risk of seeming ridiculous, I have offered this story for three reasons. First, it is an example of opposing an authority,

319

which resembles a common reason for political formation. Second, it is an example of a relationship between two parties related to that authority strengthening through the discovery of an unexpected common ground, which resembles a common basis of political formation. Third, just as an ethics for the depressed is structured by a response to one's own compulsions, so this example involves, I think, a response to another person's compulsions. I may believe that my father is responsible for the social consequences of his actions, but I do not believe that he is responsible for his atypical social intuitions, which I think motivated his conclusion that the reactions of others are fundamentally unpredictable.

In being a confrontation between individuals and not groups, it is fair to say that the example is still ethical and not yet political. But it is "proto-political" in the sense that it aims toward a different relationship that has the power to structure a community, which in this case is my family. I take the story of my conversation with my stepmother to be an example not just of my gaining more empathy for her but of a kind of consciousness raising in which I acquire solidarity with her. This situates solidarity, as a means of resistance to the power of authority, as not just a goal of political formation but a potential source of a sense of stability and thus an additional motive for the depressed. In turn, the story of my preceding confrontation with my father situates integrity as the reason for acting, in a qualitatively ambiguous situation, in accord with one's practical identity and higher values. In that case, I felt myself to respond to a kind of cosmic unfairness. Through no fault of my own, I felt, my father was behaving disgracefully before me. Through no fault of my own, he could retreat to his mother while I could not.

I thus may frame integrity as contingently good when it responds to a qualitatively ambiguous situation and incorporates higher values. This incorporates Flanagan's critique of Williams from Chapter 4, that being that alienation may be necessary for social progress and thus that to declare one's ground project categorically good, and therefore alienation from it categorically bad, may lead to either an excessively individualistic or excessively conservative rejection of social progress. I admire the heroic stand of the existentialist who acts absurdly in the name of their own freedom even when they believe their every choice to be ultimately meaningless. I think that the reconciliation of normative ethics and existentialism lies in defining this absurd stand as not absolutely required for freedom but nevertheless possible and contingently justifiable. This positions the good of integrity as a kind of existential hope.

I have attempted to show that ethics for the depressed is self-transcendent through my example of giving blood. To justify the claim that a politics for the depressed is transcendent of in-group values, I think the best example would be one in which one acting consistently with a politics for the depressed transcended their in-group's values at the risk of ostracization and thus loss of a sense of stability. I do not have such an example ready to hand. Moreover, throughout the writing of this dissertation, my values have not changed. I still believe, as I ever did, that compassion and fairness are intrinsically valuable, loyalty is intrinsically valuable but less so, and deference to authority and sanctity are just instrumentally valuable insofar as they are consistent with or support the other values. From the perspective of a conservative like Haidt, I will not have transcended my in-group values, at least so far.

But as I have asked how I should live my life and attempted to formulate an ethics for the depressed to reliably guide my answer, I have changed in my assessment of the necessity of politics. I began by simply trying to go on as an individual. Now I am convinced that I must do my part to create a world in which it is easier for us all to go on, not just through interpersonal interaction but on a grander scale, and that this effort is continuous with my effort to go on as an individual. My turn toward politics, and with it toward history and knowledge of material facts, is earnest: we need and deserve a trustworthy world.

It is plausible that I would have little of substance to say to someone who suffers from social or material precarity, not just psychological or existential. I am the son of lawyers. I feel doubt and shame at the prospect of discussing solidarity. But I respond with confidence, understood not as the absence of the awareness of risk but as the willingness to risk. Guilt would be misguided because I am not at fault for my origins. Still, fear of shame would be a reasonable response. There are contexts in which my privilege will be, I think reasonably, seen as shameful or silly. That's fine. That's just to say that some reasonable people will find me annoying. Earnestness can be annoying to reasonable people because its reach tends to exceed its grasp. I have surely overlooked, and will surely continue to overlook, what is obvious to others and so possibly waste their time. I am simply prepared to be laughed at. Well, I hope I am.[61]

I conclude that, insofar as I have lived my ethics, there is a chance that my ethics may guide others to the same turn as mine toward politics and toward solidarity as a higher source of a sense of stability, one that benefits greatly from the successful self-

---

[61] So much for confidence, huh?

rehabituation motivated and justified by the earlier stages. So long as I justify the imperative to engage in a manner guided by higher values, I think that my ethics of engagement has done, and that I have done, enough.

My story has an epilogue. Some time later, my father sent out one of his emails and mentioned, if I remember right, that he had become concerned that he was a "sociopath." He is not. No sociopath would say so. Or let his opponent take breaks.

I can no longer address my mother directly, but I can address my father and my stepmother, so I will. Dad: I love you. I'm proud to be your son. I think that what is most difficult to explain or justify is what is obvious to most people. If I have become a good philosopher, it is importantly because of you. Haru: I love you. I'm proud to be your son, too. As you have tried to speak in new ways for me, so I have tried to understand the unsaid, its meaning and its value, for you. If I have become a good philosopher, it is importantly because of you.

Dad, one last note: knowing you, it may not have occurred to you to be offended that I have represented you here with a potentially unflattering story. Even so, thanks in advance for understanding. I expect you'll want to dig up the emails that I mention in the story and epilogue and confirm what you said. Know that I chose not to confirm if my memories are correct because I intended the story to be not about the truth of these events, but about how I feel about you, which is represented in part by how I remember these events. I wanted to make that explicit.

But I will leave the rest implicit. I hope that my writing has been up to the task of making my meaning clear.

REFERENCES CITED

Alexander, Thomas. 1987. *John Dewey's Theory of Art, Experience, and Nature: The Horizons of Feeling.* SUNY Press.

American Psychiatric Association. 2000. *Diagnostic and Statistical Manual of Mental Disorders*. 4th edition. Washington, DC: Author.

———. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. 5th edition. Washington, DC: Author.

Anderson, Elizabeth. 2009. "Toward a Non-Ideal, Relational Methodology for Political Philosophy: Comments on Schwartzman's 'Challenging Liberalism.'" *Hypatia* 24 (4): 130–45.

Annas, Julia. 1995. "Prudence and Morality in Ancient and Modern Ethics." *Ethics* 105: 241–57.

Anscombe, G. E. M. 1981. *Ethics, Religion and Politics: Collected Philosophical Papers Volume III.* Oxford: Basil Blackwell.

Baier, Annette C. 1986. "Trust and Antitrust." *Ethics* 96 (2): 231–60.

———. 1987. "The Need for More than Justice." In *Justice and Care: Essential Readings in Feminist Ethics*, edited by Virginia Held, 47–58. Westview Press.

———. 2009. *Reflections on How We Live*. Oxford University Press.

Baldwin, James. 2008. "Nothing Personal." *Contributions in Black Studies* 6: 49–60.

Bernstein, Jay. 2011. "Trust: On the Real But Almost Always Unnoticed, Ever-Changing Foundation of Ethical Life." *Metaphilosophy* 42 (4): 395–416.

———. 2015. *Torture and Dignity: An Essay on Moral Injury*. University of Chicago Press.

Bernstein, Richard. 1961. "John Dewey's Metaphysics of Experience." *The Journal of Philosophy*, Vol. 58 (1).

———. 2010. *The Pragmatic Turn*. Polity Press.

Biegler, Paul. 2010. "Autonomy and Ethical Treatment in Depression." *Bioethics* 24 (4): 179–89.

Birmingham, Kevin. 2017. "'The Great Shame of Our Profession': How the humanities survive on exploitation." *The Chronicle of Higher Education*.

Bloomfield, Paul. 2009. "Review: David Wong: Natural Moralities." *Mind* 118 (469): 225–30.

Caro, Jason. 2009. "Levinas and the Palestinians." *Philosophy & Social Criticism* 36 (6): 671–84.

Cholbi, Michael. 2011. "Depression, Listlessness, and Moral Motivation." *Ratio* XXIV: 28–45.
.
Critchley, Simon. 1999. *The Ethics of Deconstruction*. Edinburgh University Press.

Crowell, Steven. 2001. "Gnostic Phenomenology: Eugen Fink and the Critique of Transcendental Reason." *The New Yearbook for Phenomenology and Phenomenological Inquiry* I: 257–77.

———. 2002. "Kantianism and Phenomenology." In *Phenomenological Approaches to Moral Philosophy: A Handbook*, edited by John J. Drummond and Lester Embree, 47–68. Springer Science+Business Media Dordrecht.

———. 2012. "Why is Ethics First Philosophy? Levinas in Phenomenological Context." *European Journal of Philosophy* 20 (4): 564–88.

———. 2013. *Normativity and Phenomenology in Husserl and Heidegger*. Cambridge University Press.

———. 2016. "Phenomenology, Meaning, and Measure." *Philosophy Today* 60 (1): 237–52.

Depraz, Natalie, Francisco J. Varela, and Pierre Vermersch. 2003. *On Becoming Aware: A Pragmatics of Experiencing*. John Benjamins.

Derrida, Jacques. 1992. "Force of Law: The 'Mystical Foundation of Authority.'" In *Deconstruction and the Possibility of Justice*, edited by Drucilla Cornell, Michael Rosenfeld and David Carlson, 3–67.

———. 2005. *On Touching—Jean-Luc Nancy*. Stanford University Press.

Dewey, John. 1930. "Qualitative Thought." In *The Essential Dewey Volume 1: Pragmatism, Education, Democracy*, edited by Larry A. Hickman and Thomas M. Alexander, 195–205. Indiana University Press.

———. 1938. *Logic: The Theory of Inquiry*. New York: Henry Holt and Company.

———. 1957. *Human Nature and Conduct: An Introduction to Social Psychology*. New York: Random House.

———. 1958. *Experience and Nature*. New York: Dover Publications.

———. 1960. *On Experience, Nature, and Freedom*. Edited by Richard Bernstein. Liberal Arts

———. 1963. *Liberalism and Social Action*. New York: Capricorn Books.

———. 2008. "Theory of Valuation." In *John Dewey: The Later Works, 1925-1953*, Vol. 13, 189–252. Southern Illinois University Press.

———. 2016. *The Public and its Problems: An Essay in Political Inquiry*. Edited by Melvin L. Rogers. Athens, Ohio: Swallow Press.

Dostoevsky, Fyodor. 2002. *The Brothers Karamazov*. Translated by Richard Pevear and Larissa Volokhonsky. New York: Farrar, Straus, and Giroux.

Dotson, Kristie. 2013. "How is this Paper Philosophy?" *Comparative Philosophy* 3 (1): 3–29.

Drummond, John J. 2002. "Aristotelianism and Phenomenology." In *Phenomenological Approaches to Moral Philosophy: A Handbook*, edited by John J. Drummond and Lester Embree, 15–46. Springer Science+Business Media Dordrecht.

———. 2009. "Feelings, Emotions, and Truly Perceiving the Valuable." *The Modern Schoolman* LXXXVI: 363–79.

———. 2013. "The Intentional Structure of Emotions." *Logical Analysis & History of Philosophy* 16 (1): 244–63.

———. 2021. "Phenomenological Method and Contemporary Ethics." *Continental Philosophy Review* 54: 123–38.

Duncan, Craig. 2016. "Review: Torture and Dignity: An Essay on Moral Injury." *Notre Dame Philosophical Reviews*.

Eisenstadt, Oona and Claire Elise Katz. 2016. "The Faceless Palestinian: A History of an Error." *Telos* 174: 9–32.

Fesmire, Steven. 2003. *John Dewey and Moral Imagination: Pragmatism in Ethics*. Indiana University Press.

Fitzpatrick, Devin. 2022. "Scheler's Phenomenology of Value as Value Pluralism." In *Max Scheler in Dialogue*, edited by Susan Gottlöber, 221–38.

FitzPatrick, William. 2012. "Book Reviews: Kitcher, Philip. *The Ethical Project*." *Ethics* 123 (1): 167–74.

Flanagan, Owen. 1991. *Varieties of Moral Personality: Ethics and Psychological Realism*. Harvard University Press.

———. *The Geography of Morals: Varieties of Moral Possibility*. 2016. New York: Oxford University Press.

Fourie, Carina. 2017. "*Who* Is Experiencing *What* Kind of Moral Distress? Distinctions for Moving from a Narrow to a Broad Definition of Moral Distress." *AMA Journal of Ethics* 19 (6): 578–84.

Frings, Manfred S. 1997. *The Mind of Max Scheler: The First Comprehensive Guide Based on the Complete Works*. Marquette University Press.

Gallagher, Shaun. 2012. "Time, Emotion, and Depression." *Emotion Review* 4 (2): 127–32.

———. 2013. "Ambiguity in the Sense of Agency." In *Decomposing the Will*, edited by Andy Clark, Julian Kiverstein, and Tillmann Vierkant. Oxford Scholarship Online.

———. 2020. *Action and Interaction*. Oxford University Press.

Gallagher, Shaun and Dylan Trigg. 2016. "Agency and Anxiety: Delusions of Control and Loss of Control in Schizophrenia and Agoraphobia." *Frontiers in Human Neuroscience* 10 (459).

Geniusas, Saulius. 2011. "William James and Edmund Husserl on the Horizontality of Experience." *Analecta Husserliana* 108.

Gouinlock, James. 1972. *John Dewey's Philosophy of Value*. Humanities Press: New York.

Graham, Jesse, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, Peter H. Ditto. 2013. "Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism." In *Advances in Experimental Social Psychology* 47: 55–130.

Greene, Joshua. 2013. *Moral Tribes: Emotion, Reaosn, and the Gap Between Us and Them*. Atlantic Books.

Hadot, Pierre. 2009. *Philosophy as a Way of Life*. Translated by Arnold Davidson. Oxford: Blackwell.

Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People are Divided by Politics and Religion*. Pantheon Books.

Hall, Cathy W. and Raymond E. Webster. 2007. "Risk Factors Among Adult Children of Alcoholics." *International Journal of Behavioral Consultation and Therapy* 3 (4): 494–511.

Hämäläinen, Nora. 2015. "Self-Help, Moral Philosophy, and the Moral Present." *Human Studies* 39 (2): 289–306.

———. 2016. *Descriptive Ethics*. New York: Palgrave Macmillan.

Heidegger, Martin. 1962. *Being and Time*. Translated by John Macquarrie and Edward Robinson. Harper: San Francisco.

Husserl, Edmund. 1983. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy: First Book.* Translated by Fred Kersten. Martinus Nijhoff Publishers.

Iurino, Kathryn and Gerard Saucier. 2020. "Testing Measurement Invariance of the Moral Foundations Questionnaire Across 27 Countries." *Assessment* 27 (2): 365–72.

Jaeggi, Rahel. 2014. *Alienation*. Translated by Frederick Neuhouser and Alan E. Smith. Edited by Frederick Neuhouser. Columbia University Press.

James, William. 2018. *Pragmatism*. Dover Publications.

———. 1983. *The Principles of Psychology*. Harvard University Press.

Johnson, Mark. 2014. *Morality for Humans: Ethical Understanding from the Perspective of Cognitive Science*. University of Chicago Press.

Kelly, Daniel and Nicolae Morar. 2014. "Against the Yuck Factor: On the Ideal Role of Disgust in Society." *Utilitas* 26 (2): 153–77.

King, Martin Luther, Jr. 1963. "Letter from Birmingham Jail." *The Atlantic Monthly* 212 (2): 78–88.

Kitcher, Philip. 2011. *The Ethical Project*. Harvard University Press.

Koopman, Colin. 2011. "Rorty's Linguistic Turn: Why (More Than) Language Matters to Philosophy." *Contemporary Pragmatism* 8 (1): 61–84.

———. 2017. "The Will, the Will to Believe, and William James: An Ethics of Freedom as Self-Transformation." *Journal of the History of Philosophy* 55 (3): 491–512.

Korb, Alex. 2015. *The Upward Spiral: Using Neuroscience to Reverse the Course of Depression, One Small Change at a Time*. New Harbinger Publications.

Landry, Christinia. 2018. "Toward a Feminist Phenomenological Ethics." In *Rethinking Feminist Phenomenology: Theoretical and Applied Perspectives*, edited by Sara Cohen Shabot and Christinia Landry, 87–102. Rowman & Littlefield International.

Law, Iain. 2009. "Motivation, Depression and Character." In *Psychiatry as Cognitive Neuroscience: Philosophical Perspectives*, edited by Matthew Broom and Lisa Bortolotti, 351–64. Oxford University Press.

Lekan, Todd. 2003. *Making Morality: Pragmatist Reconstruction in Ethical Theory*. Vanderbilt University Press.

———. 2006. "Pragmatist Metaethics: Moral Theory as a Deliberative Practice." *The Southern Journal of Philosophy* XLIV: 253–271.

Lewis, Gwyneth. 2006. *Sunbathing in the Rain: A Cheerful Book about Depression*. HarperCollins.

Mann, Bonnie. 2018. "The Difference of Feminist Philosophy: The Case of Shame." *Puncta: Journal of Critical Phenomenology* 1 (1): 41–73.

Marchetti, Sarin. n.d. "Hortatory Ethics." Unpublished draft.

Merleau-Ponty, Maurice. 1968. *The Visible and the Invisible*. Edited by Claude Lefort. Translated by Alphonso Lingis. Northwestern University Press.

Miyamoto, Musashi. 1974. *A Book of Five Rings: The Classic Guide to Strategy*. The Overlook Press.

Mouffe, Chantal. 1996. *Deconstruction and Pragmatism: Simon Critchley, Jacques Derrida, Ernesto Laclau and Richard Rorty*. Routledge.

Nietzsche, Friedrich. 1966. *Beyond Good and Evil: Prelude to a Philosophy of the Future*. Translated by Walter Arnold Kaufmann. New York: Vintage Books.

Noggle, Robert. 2016. "Belief, Quasi-Belief, and Obsessive-Compulsive Disorder." *Philosophical Psychology*.

Pappas, Gregory Fernando. 2008. *John Dewey's Ethics: Democracy as Experience*. Indiana University Press.

Petitmengin, Claire. 2006. "Describing One's Subjective Experience in the Second Person: An Interview Method for the Science of Consciousness." *Phenomenology and Cognitive Science* 5: 229–69.

Prinz, Jesse. 2004. "Embodied Emotions." In *Thinking about Feeling: Contemporary Philosophers on Emotions*, edited by R. C. Solomon, 44–58. Oxford University Press.

Ratcliffe, Matthew. 2008. *Feelings of Being: Phenomenology, Psychiatry and the Sense of Reality*. Oxford University Press.

———. 2015. *Experiences of Depression: A Study in Phenomenology*. Oxford University Press.

———. 2020. "Existential Feelings." In *The Routledge Handbook of Phenomenology of Emotion*, edited by Thomas Szanto and Hilge Landweer, 249–60. Routledge.

Ratcliffe, Matthew, Mark Ruddell, and Benedict Smith. 2014. "What is a 'Sense of Foreshortened Future?' A Phenomenology Study of Trauma, Trust, and Time." *Frontiers in Psychology* 5: 1026.

Redding, Paul. 2020. "Robert Pippin's Hegel as an Analytically Approachable Philosopher." *Australasian Philosophical Review* 2 (4): 355–64.

Roberts, John Russell. 2001. "Mental Illness, Motivation and Moral Commitment." *Philosophical Quarterly* 51 (202): 41–59.

Rorty, Richard. 2004. "Philosophy-envy." *Daedalus* 133 (4): 18–24.

Ruprecht, Louis A., Jr. 2017. "Cornel West and the Tragedy at the Heart of North American Pragmatism: A Retrospective Look at The American Evasion of Philosophy." *American Journal of Theology & Philosophy* 38 (2-3): 179–200.

Saarinen, Jussi. 2018. "A Critical Examination of Existential Feeling." *Phenomenology and the Cognitive Sciences* 17 (2): 363–74.

Sager, Alex. 2014. "Normative Ethics after Pragmatic Naturalism." *Metaphilosophy* 45 (3): 422–440.

Scheffler, Samuel. 1982. *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. Clarendon Press.

———. 1992. *Human Morality*. Oxford University Press.

Scheler, Max. 1954. *The Nature of Sympathy*. London: Routledge & Kegan Paul.

———. 1973a. *Formalism in Ethics and Non-Formal Ethics of Values: A New Attempt toward the Foundation of an Ethical Personalism*. Translated by Manfred S. Frings and Roger L. Funk. Northwestern University Press.

———. 1973b. *Selected Philosophical Essays*. Translated by David R. Lachterman. Northwestern University Press.

———. 1972. *On the Eternal in Man*. Translated by Bernard Noble. Archon Books.

———. 1998. *Ressentiment*. Translated by Lewis B. Coser and William A. Holdheim. Marquette University Press.

Schwartz, Shalom H. 2012. "An Overview of the Schwartz Theory of Basic Values." *Online Readings in Psychology and Culture* 2 (1).

Schwitzgebel, Eric and Joshua Rust. 2014. "The Moral Behavior of Ethics Professors: Relationships Among Self-Reported Behavior, Expressed Normative Attitude, and Directly Observed Behavior." *Philosophical Psychology* 27 (3): 293–327.

Sigrist, Michael J. 2019. "Why Aren't Ethicists More Ethical?" *Blog of the APA*.

Smith, Benedict. 2013. "Depression and Motivation." *Phenomenology and the Cognitive Sciences* 12 (4): 615–35.

Smith, Michael. 1994. *The Moral Problem*. Blackwell Publishing.

Smith, William Hosner. 2012. *The Phenomenology of Moral Normativity*. London: Routledge.

Solomon, Andrew. 1998. "Anatomy of Melancholy." *The New Yorker*.

Stawarska, Beata. 2018. "Subject and Structure in Feminist Phenomenology: Re-reading Beauvoir with Butler." In *Rethinking Feminist Phenomenology: Theoretical and Applied Perspectives*, edited by Sara Cohen Shabot and Christinia Landry, 13–32. Rowman & Littlefield International.

Stein, Edith. 1989. *On the Problem of Empathy*. Third edition. Translated by Waltraut Stein. Washington, D.C.: ICS Publications.

———. 2000. *Philosophy of Psychology and the Humanities*. Edited by Marianne Sawicki. Translated by Mary Catharine Baseheart and Marianne Sawicki. Washington, D.C.: ICS Publications.

Stocker, Michael. 1979. "Desiring the Bad: An Essay in Moral Psychology." *Journal of Philosophy* 76 (12): 738–53.

Strawson, P. F. 1962. "Freedom and Resentment." *Proceedings of the British Academy* 48: 1–25.

Swartzer, Steven. 2015. "Humean Externalism and the Argument from Depression." *Journal of Ethics and Social Philosophy* 9 (2): 1–16.

———. 2018. "A Challenge for Humean Externalism." *Philosophical Studies* 175 (1): 23–44.

Tiberius, Valerie. 2015. *Moral Psychology: A Contemporary Introduction*. London: Routledge.

Tiles, J. E. 2004. "Review of John Dewey and Moral Imagination: Pragmatism in Ethics, by Steven Fesmire." *Transactions of the Charles S. Peirce Society* 40 (2).

———. 2009. "Review of Gregory Fernando Pappas, John Dewey's Ethics: Democracy as Experience." *Ethics* 119 (2).

Vasiliou, Iakovos. 2016. *Moral Motivation: A History*. Oxford University Press.

West, Cornel. 1989. *The American Evasion of Philosophy*. Palgrave Macmillan.

Williams, Bernard. 1973. *Problems of the Self: Philosophical Papers 1956-1972*. Cambridge University Press.

———. 1981. *Moral Luck: Philosophical Papers, 1973-1980*. Cambridge University Press.

———. 2006. *Ethics and the Limits of Philosophy*. Routledge.

———. 2015. *Morality: An Introduction to Ethics*. Cambridge University Press.

Wolfe, Joel. 2012. "Does Pragmatism Have a Theory of Power?" *European Journal of Pragmatism and American Philosophy* IV (1).

Wong, David. 2006. *Natural Moralities: A Defense of Pluralistic Relativism*. Oxford University Press.

Yong, Ed. 2021. "Why Health-Care Workers are Quitting in Droves." *The Atlantic*.