Bad Feelings, Feeling Badly

Erica Bigelow

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2026

Reading Committee:

Carina Fourie, Chair

Sara Goering

Carole Lee

Gaile Polhhaus, Jr.

Program Authorized to Offer Degree:

Department of Philosophy

University of Washington

Abstract

Bad Feelings, Feeling Badly

Erica Bigelow

Chair of the Supervisory Committee:

Carina Fourie

Department of Philosophy

"Emotions have been constructed as the other of thought, rationality, and reason, and they have also been projected onto the bodies of gendered and racial others," writes Sara Ahmed in discussing the politics of fear (2003, 381). Ahmed is one node in the lineage of feminist theorists whose work calls for a new attention to emotion as a fundamental dimension of our inner and outer lives, particularly as one that's shaped by interpersonal and structural injustices. In keeping with this tradition, my dissertation asks *what we owe to others' emotions* by considering the emotions as normatively evaluable. Chapter 1 studies the weaponization of emotions through a critical disability studies lens. Chapter 2 presents the nascent affective injustice literature as an attempt to identify wrongdoings that target the emotions, and points to some shortcomings of this literature. In chapter 3, I look at a particular species of emotional wronging that relies on the weaponization not of any particular emotion, but of sociocultural norms governing our interactions with others' emotions. Chapter 4 explores the prevalence of emotional wronging in online communities. In chapter 5, I propose *attentive care* – a blend of traditional care ethics with Iris Murdoch's writings on just attention – as a way of making sense of our obligations toward others in their capacities as *feeling* beings.

ACKNOWLEDGMENTS

This dissertation was a collaborative effort and was made possibly only by the support of my committee members: my advisor and chair Carina Fourie, my unofficial second advisor Sara Goering, Carole Lee, Gaile Pohlhaus, and Tim Brown. I am also indebted to my undergraduate advisors – Megan Mitchell, Andrea Opitz, and Laura Scales – for allowing a younger version of me to see herself as a scholar. The UW Philosophy faculty and staff, including Carla Saulter, Jana Mohr Lone, Karen Emmerman, and Paul Franco also provided critical support during this undertaking. Moreover, the questionably-named Boltzmann writing group (Amelia, José, Josh, Rose, Shawn, and Erin) is where many of the finishing touches of this dissertation were put in place. Over the course of writing this dissertation, I received additional financial support from the Philosophy Learning and Teaching Organization (PLATO), the Simpson Center for the Humanities, the Gerler family, the Marc Sanders Foundation, the Emotion & Society Lab, and the Institute for Critical Disability Studies at Western Washington University.

I became an aunt during my first quarter of graduate school. My nephews, Charles Takeru Bigelow and Henry David Gaudreau, remain the most influential philosophers in my life. Sage Feltus and Sosseh Assaturian, my best friends, I love you dearly. Special thanks go to my siblings – Shel, Brian, Dan, Tyler, and Amy. Living thousands of miles away from Massachusetts somehow made me closer than ever with my parents, and I will forever be grateful for their love and support.

To everyone else who made this possible – Kit and Kai Clement-Milanovich, Leah Davis, Kade Cicchella, Natalie Dorfman, Amelia Wirts, Mikala MacLennan, Jesus Raya, Julia Pelger, Elina Castellano-Tejada, Rose and Audrey Sanchez, Gabbi Courtenay, Tobias Romein, Randy Rodriguez, my Miami (*not Florida*) people – 'thank you' will never be enough. Lastly, to my COVID cohort mates Bri High and Nicolai Wohns – we're a long way from our Zoom hangouts.

DEDICATION

This dissertation is dedicated to Amy Aiko Kawazoe and Louis George Bigelow. You should be here.

Introduction: Feeling Responsibly

It is not uncommon in online spaces, particularly those oriented toward social justice, to aver that we 'don't owe anyone anything.' This idea is often invoked as shorthand to inure us against manipulation and abuse, or against feeling guilty for *refusing* to be so manipulated or abused. Moreover, it takes the responsibility for a wide range of injustices off of individuals' shoulders by limiting their obligations to intervene; though the state may owe someone something by virtue of their belonging to its people, or parties to a contract may owe one another whatever is dictated by that contract, I as an individual am (by default, qua individual) exempt from such obligations. However well-intended (enticing, even) the idea that we owe others nothing may be, in many of its applications I find it misguided at best, and harmful and anti-social at worst, as it grants us tacit permission to eschew those obligations that we *do* have. Phrases like it should be subject to stricter philosophical scrutiny, moreover, insofar as they convey *oughts*; even if it is presented as a descriptive claim about obligations that we do or do not have, it is action-guiding. It tells us how to react to others, and how to expect others to react to us. The premise that I begin this dissertation from, then, is a response to a question about oughts: I contend that we owe others (and they in turn owe us) quite a considerable amount, as it is only on the basis of this reciprocity that any remotely democratic system can thrive. I am especially interested in oughts (obligations) that arise within our affective lives, particularly as the notion of *not owing anyone anything* in terms of emotional investment suggests that any normative analysis is ill-suited for a conversation about emotions.

Sara Ahmed writes in a 2003 essay that "[e]motions have been constructed as the other of thought, rationality, and reason," pointing to the historical tradition — one that has been unsettled in recent decades but was once especially strong within academic philosophy because of the discipline's appreciation for logic and rationality — of rendering emotions a- or il-logical[1]

[1] Jaggar (1989) writes: "From Plato until the present, with a few notable exceptions, reason rather than emotion has been regarded as the indispensable faculty for acquiring knowledge" (151). She adds, too, that

(381). Ahmed is one node in a feminist lineage whose work renews attention to emotion as a fundamental dimension of our inner and outer lives. In what follows I continue this tradition by studying the myriad ways that emotions are related to harm. The different dimensions of this relationship are explored in more depth below. Consequently, emotions can be the subjects of 'oughts' like those mentioned in the previous paragraph That is, rather than merely acknowledging the emotional fallout of wrongdoings, or of other non-emotional harms, I look to emotions as sometimes constituting the *stuff* of wrongdoings, as themselves wrongdoings. I begin, then, from the premise that emotions can be normatively evaluated, and that they can even be *wrong* insofar as they can constitute *wrongs*. Emotions can be, in other words, unethical; there are certain ways that we ought not to feel.

I want to be clear about terminology, as there are a few language choices throughout this dissertation that might otherwise seem ambiguous. The following chapters — the second in particular with its focus on the affective injustice literature — use the word (in)justice frequently in talking about the different wrongs that individuals can commit against others in their capacities as affective beings. In so doing, though, I do not want to limit (or even focus) the conversation to (on) large-scale structural injustices, those injustices that moral and political philosophers often mean by the term. Instead, I am concerned primarily with the ways that these injustices — when taken for granted as such — are instantiated between or among individuals. The 'injustices' I discuss, then, need not actually be injustices per se, if we understand injustice as having some necessary relationship to the usual referents of social (in)justice; they are more akin to harms or what DuFord terms "substantive conflicts [:] conflicts over elements of the world that affect our lives in serious and thoroughgoing ways" (2022, 7). Such conflicts, DuFord writes, need not get in the way or foreclose the possibility of solidarity; they do not rise to the level of individuals (e.g.) dehumanizing one another in the ways that

"the western [philosophical] tradition has not seen everyone as equally emotional" (ibid, 163). Thus, not everyone is seeing as having an equal claim to or stake in philosophical knowledge-building projects.

injustice involves. Taking these conflicts seriously is important because it allows us to reckon seriously with intragroup conflict without defaulting to the view that such conflicts are themselves injustices, as doing so might inadvertently 'water down' injustice, making it a catch-all term for cases where people butt heads. The contents of such conflicts are no doubt important, and they may certainly include references to injustice, but it is also important that we be able to discuss them without automatically painting one party as (for instance) oppressive (as we do, often rightfully, when talking about injustice). Moreover, thinking in these terms (that is, in terms of interpersonal conflict or harm rather than injustice) empowers a theory to address conflicts that arise among relative equals, individuals who are similarly situated with respect to whatever is at issue. Another way of framing this appears in the terms of Sarah Schulman's distinction between conflict and abuse, and her indictment of habitually *overstating harm* (2016, ch. 3). Abuse exists, just as injustice exists. But not all conflict is abusive, just as not all acts of harm are unjust. Thus, when I use the word 'injustice' and its variants throughout this dissertation, I am using it as a term that often overlaps with, but is not coextensive with, interpersonal (i.e., smaller scale) conflict.

Also worth noting is that much of what follows uses the words 'affect' and 'emotion' as interchangeable, though some insist that the two have different referents. This project began through my interest in the affective injustice literature as an alternative to epistemic injustice, so I initially anticipated it requiring a deep dive into affect theory. Ultimately, I have found that those who write about affective injustice do so in a way that treats 'affective' and 'emotional' as synonyms, so I have chosen to follow their leads.

Emotions show up throughout this dissertation in myriad ways, and the terms I use throughout are contested. Even the division, the valence-ing, of emotions into positive and negative ones – irrespective of their normative value, i.e., whether they are morally laudable – is more complex than it seems. I am primarily concerned with the intersection between emotions and wrongfulness, but this wrongfulness does not have a single, simple referent. Among the ways emotions appear in what follows are: as wrongful in themselves, as brought about by some wrong, as things that we can act better or worse toward, and as things that are wrongfully disincentivized (as in some forms of apathy). In short, I am not concerned with merely one intersection of wrongfulness or wronging and emotions; the two concepts can interact in a wide range of ways.

To reiterate: all these dimensions of wronging and emotions are distinct from an emotion's valence. That is, the dynamics I am studying can happen with typically negative emotions (anger, sadness, pity) and with typically positive ones (glee, awe, contentment). This positivity and negativity have to do with an emotion's placement on a general, colloquially understood good-to-bad scale. When discussing an emotion as positive or negative, I will use the broad (if vague) criteria according to which we typically classify emotions as such: "the mundane…contrast between good and bad" (Solomon and Stone 2002, 420). This hinges, too, on approval and disapproval; a feeling we approve of is good (positive) and one we disapprove of is bad (negative). I find this distinction at least more helpful than linking the goodness of a particular emotion to (for example) its tendency to produce pleasure, as this becomes more complicated when we try to ascertain why pleasure is (often thought of as) a positive emotion. Speaking about the tendency to produce pleasure as the defining factor in whether an emotion is positive or negative might, too, be too consequentialist a view (see Solomon and Stone 2002, 423).

When I speak, toward the end of this dissertation, of attentive care, I do so as someone concerned with lessening the likelihood of harm, and of making our interactions with others more adherent to what I refer to throughout this dissertation as an affective ethics, drawn out further in the fifth and final chapter. I do not purport to be providing a complete theory of justice (since injustice is not all that I am concerned with; affective ethics often calls for interventions more mundane and localized than *doing justice*) nor am I suggesting that harm will be impossible under an attitude of attentive care. Instead, I am suggesting a moral disposition, a moral outlook, that helps us to see better, and so to see (others, their emotions, etc.) in ways that are less likely to be in conflict with how they want to be seen. This is a disposition that can be taken on not only in moments of extreme conflict motivated by social injustice, but even when we find ourselves butting heads with our all-things-considered equals.

Overview By Chapter

The following project consists of five chapters, encompassing both theoretical and applied work. I begin in Chapter 1 with examples that, owing to their grounding in social injustice, might make harmful affective practices more obvious by inviting us to call into question the normative status of particular emotions, namely those that serve to maintain anti-disability prejudice (or ableism). The three example affects here are pity, fear, and awe. Though none of these emotions is harmful in isolation — it would be inappropriate to say that, for instance, *all fear causes harm and should thus be habituated against* — when directed at disabled people on the basis of disability they can come to convey (and even enforce) particular harmful ideas about disability. On this basis I begin to demonstrate the need for an affective ethics, because we ought to be able to argue that such emotions are (when used in service of ableism) *unethical*.

Chapter 2 provides an overview and a critique of the nascent affective injustice literature, grounded in 2018 papers by Amia Srinivasan and Shiloh Whitney. My concerns with the affective injustice literature are somewhat paradoxical: I worry that it both lacks an account of when an affective harm is escalated to the level of injustice, and that it simultaneously relies too heavily on social injustice as its basis, obscuring the more mundane ways that individuals wrong one another. Accordingly, the affective injustice literature is unable to account for more complex cases, including those that either (1) remain unspoken and thus do not find cause harms via their uptake, or (2) happen among equals rather than being rooted in (what Srinivasan calls) "first-order injustice," which seems to refer to identity-based structural mistreatment (2018, 135). So, in some sense the affective injustice literature's target is too narrow, but in others it is too broad. Thus, though the label 'affective (in)justice' serves its authors' purposes quite well, harms like those mentioned in Chapter 1 and ones that occur between or among similarly situated individuals are left out of the picture.

In Chapter 3, I articulate another case that underscores the need for an affective ethics, and which obviates that some of the affective injustice literature is too narrow. Much of the affective injustice literature at the center of Chapter 2 makes implicit references to emotion norms: social norms dictating what emotions are appropriate in which situations. I provide a brief account of emotion norms, using as my paradigm cases where one party's anger is rhetorically transformed into Madness as an example where social norms and psychiatrization merge. I then describe another set of norms relevant to my affective ethics: norms pertaining not to emotions themselves, but to the nexus of behaviors surrounding how we speak about and behave toward emotional claims, using the notion of 'therapy-speak' as an example. I term these emotion-responsive norms, or 'responsive norms' for short.

Following Chapter 3's norm-laden discussion, Chapter 4 explores the maintenance of such norms in a setting that is increasingly becoming the space where emotions are mediated: cyberspace, particularly social media platforms. I study such platforms in terms of three infrastructural levels — usage patterns, on-platform relationships among users, and content — arguing that on each of these levels, apathy toward others' emotions is normalized and incentivized. The phenomena identified in this chapter thus demonstrate how background structures of discourse — as in the literal digital infrastructure through which discourse is mediated — can incline us to respond to others in harmful ways. As we see in this chapter, these do not necessarily mirror patterns of social injustice, particularly given that social media platforms often give their users relatively equal 'standing'.

The crux of my positive argument comes in Chapter 5. In this chapter, I propose *attentive care* as an ameliorative affective disposition that, when adopted, provides ethical guidance for our interactions with others' emotions. Attentive care combines the core tenets of feminist care ethics with Iris Murdoch's morality centered on *attending* to others. I conclude this chapter, and the dissertation itself, by applying attentive care ethics to the examples discussed in chapters 1, 3, and 4, making clears its utility as an affective ethics. What makes attentive care ethics such an appealing proposal is its malleability; because it does not prescribe any precise action beyond a way of seeing, it is relevant in virtually all aspects of life, including those where structural social injustice plays little to no role.

Chapter 1 - Bad Feelings: The Emotions of the Ableist Imaginary

Introduction

What does it mean to have *bad feelings*? In the colloquial sense as described in the introduction to this dissertation, bad feelings are “negative” emotions; to have a bad feeling about something might mean to have a sense of foreboding, one which we generally perceive as bad in part because it is unpleasant[2]. But even though these emotions *feel bad*, I am not convinced that this makes them *bad feelings*, at least in the sense that I am concerned with. Some, but not all, negative emotions are, strictly speaking, bad feelings. They are feelings that are bad, not just because they feel bad, but because they express or communicate some condemnatory attitude. Moreover, feeling bad is not sufficient for a *feeling* to *be bad*; some people might find delight in what I am considering bad feelings.

Consider this example Sara Ahmed draws from Frantz Fanon’s *Black Skin, White Masks*:

> ‘Mama, see the Negro! I’m frightened! Frightened! Frightened!’ Now they were beginning to be afraid of me. I made up my mind to laugh myself to tears, but laughter had become impossible (Ahmed 2014, 62; Fanon 1982, 111-12).

Here, Fanon is describing his being declared the object of a white child’s fear. Ahmed studies this fear as “re-establish[ing] distance between bodies” in proximity to one another; the child’s fear makes it clear that they view themselves as fundamentally different than, apart from, a Black man, even though they may share physical space (2014, 63; Ahmed 2004). Ahmed notes that the very appearance of fear is constitutive of an “affective politics” that presumes a relationship between Blackness and threat (ibid). This is a politicizing of affect or emotion

[2] To reiterate an earlier claim: I do not find pleasure or pleasantness alone a strong criterion by which to evaluate emotions.

because it is both responsive to political conditions and because it serves to maintain those conditions. What I want to expand upon, though, is that *fear* as a first-line response to otherness can not only be or facilitate poor politics but it can be morally *wrong*. In this first chapter, I sketch the outlines of an affective ethics by looking at *un*ethical emotions, focusing on the ways that emotions can be the tools of ableism or anti-disability prejudice. Insofar as they are morally wrong, I contend, these emotions make an affective ethics necessary.

There are two levels of analysis at play here: first, I consider what Thomas (1999) calls *psycho-emotional disablism*, which entails looking at the ways that experiencing ableism produces negative[3] emotional outcomes *for disabled people*. Second, I look at the ways that ableism can be manifested in its perpetrators' attitudes, which themselves are used to harm disabled people. In the latter respect, I study three specific affects or emotions within what I call (following exemplars including West & Johnson [2024]) the *ableist imaginary* — the system of ideas, assumptions, ideals, and ignorances (deliberate and otherwise) that uphold anti-disability discrimination. These affects are pity, fear, and awe. I conclude the chapter by (re)defining the phenomena described in this section about the ableist imaginary as cases of emotional *weaponization*. In other words, the three emotions identified — pity, fear, and awe — are, I argue, weaponized in service of the ableist imaginary. I thus invoke a normative evaluation of emotions, suggesting that in such weaponized uses these emotions themselves can be wrongful, rather than needing to look to the wrongful effects of (other acts of) ableism as Thomas' work on psycho-emotional disablism does. These are cases where unethical affects map neatly onto other social prejudices, but as I show in later chapters, this is not necessary. In other words, unethical

[3] When I write in this chapter about negative emotions, I am using the value judgment in the colloquial sense. The negative emotions, the *bad feelings*, I write about here might align with those that Sianne Ngai identifies as ugly feelings, ones "that are organized by trajectories of repulsion rather than attraction, by phobic strivings 'away from' rather than philic strivings 'toward'" (2005, 11). By this, I mean they are emotions that often represent a lashing-out at (or distancing from) the world around us, in addition to incentivizing others to remain distance from us. The value judgment is not really much of a value judgment at all, then, and is more representative of "aversion, exclusion, and of course negation" (ibid, 12).

emotional praxes can arise not only in tandem with long-recognized social injustices (like ableism), but also between individuals, even individuals who are situated relatively equally.

Bad Feelings: On Psycho-Emotional Disablism

Ableism, Disablism, and the Problem of Pain

This subsection clarifies the target of analysis in this chapter; what does it mean to describe the emotions of *ableism*, or those that exist within the *ableist* imaginary? To begin, the terms 'ableism' and 'disablism' tend to be used interchangeably, but the two technically target different phenomena, or at least examine the same phenomena from different perspectives. In the next section I take Carol Thomas' writing about psycho-emotional *disablism* to be just as relevant to conversations about *ableism* because of the ways the words are colloquially used. I begin this section, though, by discussing the ways that these concepts are sometimes defined distinctly from one another.

Disablism, writes Dan Goodley (2014), is a term describing social practices that "threaten to exclude, eradicate and neutralise those individuals, bodies, minds and community practices that fail to fit the capitalist imperative" by virtue of their disability, actual or supposed[4] (xi). Even more directly, disablism looks like disabled persons' exclusion (or *marginalization*) from participating in certain social rituals (employment, education, family-building, and so forth) (Goodley 2014, 19; Miller et al. 2004, 19). This process of exclusion is expedited when (e.g.) social programs that facilitate disabled people's inclusion in society are defunded or eliminated altogether; marginalization becomes the default when modes of inclusion are removed. As a result, disabled people are prevented from being able "to participate as full members of society," insofar as these rites are constitutive of social membership (Miller et al. 2004, 22).

[4] That is, even nondisabled people can be subject to disablism if they are perceived (and consequently treated) as disabled.

Ableism, on the other hand, refers to "the favoritism for certain abilities" that some social practices exhibit (Goodley 2014, 22). Ableism thus focuses not on the exclusion of one group (disabled people), but on the sole *in*clusion — the championing, even — of another group (in this case, nondisabled or able-bodied people). Where disablism is a preference (to put it lightly) *against*, ableism is a preference (again, an understatement) *for*. The two can be understood as the positive and negative sides of the same coin, given that their referents are constructed in opposition to one another. The distinction between ableism and disablism is often muddled, and the two are treated as synonyms in much public discourse about disability – and they will be treated as synonymous throughout this dissertation. Goodley identifies "ableism's tacit, hidden, masked, accepted, hegemonic privileging of a distinct idea of humanity" as "the hidden referent of disablism," suggesting that the two exist dialectically (2014, 26). That is, every preference-against entails a preference-for and vice versa. Moreover, Michelle Nario-Redmond's definition of ableism (not disablism) as "prejudice and discrimination toward individuals simply because they are classified as disabled" recognizes that ableism would not exist if the able-bodied did not stand out in contrast to some other group — in this case, disabled people (2019, 5). Thus, while I want to recognize that some within the critical disability studies literature treat the two (ableism and disablism) as distinct, this distinction is often quite blurry and committing to it is unnecessary for the purposes of this dissertation. Therefore, I have chosen to treat 'disablism' and 'ableism' as synonyms.

It is important to maintain here that when I talk about ableism, I am talking about a form of prejudice against disabled *people*. This does not necessarily entail any commitment to the normative value of disability *itself*, that is, disability if we were able to conceive of it as wholly separate from disabled people and their social standing. To illustrate this, consider what some have termed the problem of pain[5], which refers to the difficulty that critical disability studies

[5] This phrase is used by Joel Reynolds (2022, 5). I describe Reynolds' discussion of the ableist conflation of disability and pain at more length below. Margaret Price (2015) suggests: "that DS [disability studies]

seems to have with recognizing the negative aspects of disability (including, in many cases, pain and suffering) without endorsing ableist or even eugenicist views espousing the outright eradication of disability. That is, if a disability primarily manifests through intense chronic pain, it seems difficult to maintain any position other than the belief that *that* disability is something individuals would be better off without. One way of framing the problem of pain is as a slippage between talking about disabled people and talking about disability itself. The harms that I detail in the following sections might seem accidental or secondary to the harms of pain. Pain (as a byproduct or symptom of many disabilities) is easier to make sense of because it strikes us as obviously negative: *sometimes, being disabled feels bad, and sometimes these bad feelings would persist even in the most access-oriented, anti-ableist, pro-inclusion society imaginable.* This is especially true of conditions of which *persistent* pain is a symptom. Reckoning with this question — this *problem* — of pain has remained a problem for disability studies scholars, particularly those who are committed to the strictest versions of the social model of disability according to which the chief problems of disability can be rectified through socio-structural change. I don't intend to solve the problem of pain here. What I want to stress, though, is that we can and should critique ableist views that disparage *disabled people* without committing ourselves to the idea that disability is itself *always* positive or even neutral.

Here a distinction recalled by Watermeyer (2017) is useful. He writes that the "feelings of grief" wrought by disabled identity owing to "the socially caused losses of disablism" need to be separated from accounts that paint disability writ large as tragic privations or deprivations of some fundamental capacities (143; 152). Watermeyer analogizes disabled grief to Cheng's *racial melancholia*, which draws out the link between racialization (and minoritization) and grief (Cheng 2000). This is not to say that being a person of color is enough to warrant grief on its own, but that the social structures of racialization, racial marginalization, and so on often render

needs to pay more attention to pain in the world of disability" (274).

racial minority status as grief worthy. Racism, in other words, often makes being a person of color unpleasant. With respect to disability, Watermeyer writes:

> first, describing a melancholic relational environment does not preclude alternative responses to social injunction, which can be self-affirming and subversive. Instead, it delineates the emotional forces within which such strategies must develop. We need an analysis which avoids both the essentializing diagnosis of "self-hatred" and the dangerous view, based on a liberal denial of difference, that othering leaves no lasting emotional impressions (154).

In other words, to avoid the problem of pain, we need to recognize that disability *can* be painful, but that it categorically is made up of more than just pain. To reiterate, the theoretical distinction between disability and disablism cannot be overstated, and we should allow for study of the emotional harms of experiencing disablism without leaning too far into tropes that paint disability as consisting solely of such harms.

Of course, physiological symptoms like pain can cause adverse emotional effects; Carol Thomas dubs these "psycho-emotional impairment effects," drawing on the social model of disability's impairment/disability distinction (1999, 47). Disabled people can and often do experience pain, fatigue, nausea, discomfort, and "somatically engendered anxiety or disorientation," none of which are pleasant experiences and all of which can cause or exacerbate psychological distress (ibid). We can recognize the negativity of pain and discomfort, though, without giving in to the "ableist conflation [which] leads people to assume that many, if not most, disabilities involve constitutive or even consuming[6] pain" (Reynolds 2022, 57). The flip side of this conflation, though, is that we should be similarly wary of painting over the unpleasant aspects of disability; we cannot deny that sometimes, a person's disabilities do involve intense and prolonged pain for that person. Throughout the rest of this chapter, the negative emotions I'm concerned with are those that manifest ableism, and not the emotions of

---

[6] Constitutive and consuming pain, respectively, refer to pain that "substantively and definitively enters into one's life project" and that which is "so intense that one's experience primarily or solely consist[s]" of it (Reynolds 2022, 52; 54). In more colloquial terms, these are types of severe pain that inhibit people's daily lives.

disability itself. Like Thomas, I want to set aside these emotional 'impairment effects' and focus instead on the ways that emotions shape experiences of ableism, rather than of disability simpliciter. Thus, I proceed by discussing two dimensions of the affective experience of ableism: the phenomenon of psycho-emotional disablism, or the affective impacts that ableist restrictions have on disabled persons, and (in the following section) the affects that exist within the ableist imaginary and thus motivate and sustain ableism.

Psycho-Emotional Disablism

Perhaps the most explicit treatment of the emotional dimensions of ableism comes in Thomas' *Female Forms: Experiencing and understanding Disability* (1999). In this text, Thomas references the "psycho-emotional dimensions of disablism,"[7] which she takes to mean those "dimensions of socially imposed restrictions…which operate to shape personal identity, subjectivity or the landscape of our interior worlds - and work along psychological and emotional pathways" (1999, 46). Henceforth, I will use the shorthand 'PED' for psycho-emotional disablism. Thomas takes PED to consist not of the emotional residue of impairment effects — the ways that being in pain can lead to emotional distress, for instance — but of the "socio-cultural processes which generate negative attitudes about impairment and disability, and sustain prejudicial meanings, ideas, discourses, images and stereotypes" (ibid, 56). Donna Reeve further divides PED into two categories: direct and indirect. Direct PED is interpersonal; it is about how strangers respond to disabled persons, "either by saying something inappropriate or by avoiding the disabled person altogether," whereas indirect PED "arise[s] from the experience of *structural* disablism" (2019, 104; 106, emphasis added).

Among examples of indirect PED are the emotional costs associated with living within 'crip spacetime' as Margaret Price (2024) describes it. Put most simply, crip spacetime can be defined as "a material-discursive reality experienced by disabled people" wherein the choice to use the

[7] Thomas defines disablism as "a particular form of unequal social relationship which manifests itself through exclusionary and oppressive practices…at the interpersonal, organizational, cultural and socio-structural levels" (1999, 40).

descriptor 'crip' rather than 'disabled' can be understood as a political and methodological choice highlighting disability's "disruptive potential" (Price 2024, 7; 11). It is, in other words, a way of capturing the spatiotemporal realities that disabled people face, which may not be apparent to the nondisabled. Seeking accommodation is emotionally difficult, as is the "constant impression-management work" that disabled people who are aware of ableist stereotypes and lines of thinking must perform (Price 2024, 117). This difficulty is what Price labels the *emotional cost* of living within and according to crip spacetime. This cost might manifest as "tension, weariness, rage, or grief" at trying to secure accommodation, or it might look like what Annika Konrad terms "access fatigue," the constant labor required to make spaces accessible and/or make others comfortable with disabled people's presence (Price 2024, 37; Konrad 2021). Such costs — measured in terms of time, effort, or even money — are harmful to disabled people.

There is a hesitation within disability studies, writes Reeve in an earlier piece, to talk about the negative emotions associated with disability and with ableism because doing so might reify a "psychopathology of disability" (2006, 94). Such a psychopathologizing account would essentially recreate the medical or individualistic model of disability by suggesting that disability is best understood by looking at individuals' psyches or minds. It would shift the focus of conversations to individual psychological ailments. Reeve and Thomas alike resist this by looking at how these bad feelings are imposed, rather than studying them as symptoms or impairments in themselves. This further underpins the shift, represented here, of discussing the affective dimensions of ableism, rather than those of disability[8] (or impairment) itself.

Thomas uses disabled women's personal narratives to make the prevalence of PED clear. She cites one woman whose disability led her to feel "self-doubt and sadness," particularly in the face of feeling rejected by her former friends, and another's testimony of developing "negative self-image and poor self-esteem" owing to the ways that people responded to her disability (Thomas

[8] Following a social model distinction between impairment and disability means that the affective dimensions of disability are more or less the same as the affective dimensions of ableism.

1999, 49; 50). The responses that disabled people receive from others often communicate (explicitly or implicitly) messages about the social status of disability. These messages then get internalized by disabled persons and 'reflected' via their negative affective responses. Moreover, there can be anticipatory negative emotions associated with ableism too; the mere *fear* of mistreatment weighs heavily on those impacted by ableism.

In *Staring*, Rosemarie Garland-Thomson reflects on the stare as a response to disability. She describes certain instances of staring — especially those in which the stared-at is visibly disabled — as manifestations of the starer being *stunned* or *alienated* by the presence of disability (2009, 20). In other words, the starer is recognizing the stared-at as something unexpected. There is a social hierarchy inherent in staring: "we stare at what visually leaps toward us…while trying not to be too eye-catching ourselves," because to welcome others' stares upon us would be to internalize that we are *other* to them (Garland-Thomson 2009, 34). When disabled people become the objects of stares, as they often are (Garland-Thomson describes the disabled body as "the exemplary form of the unforeseen" [ibid, 38]), in addition to any base-level discomfort with having their actions monitored so closely, they have their otherness, their distance from normativity, reinforced. This, quite simply, *does not feel good*. The "pain of exclusion" can take different forms as evidenced by the table reproduced below, but it seems to convey a central message: "that it is normal and natural for disabled young people to be *absent*" (Murray 2006, 38; 36). This is not exclusive to staring behaviors; it is true of any "experience of being excluded from physical environments," and so applies to the sorts of physical infrastructure that more traditional readings of the social model point to as disabling too (Reeve 2006, 97).

Exclusion	**Inclusion**
Useless, unwanted, depressed, lonely, isolated	Warm, secure, anything is possible
Sad, angry, frustrated	Important, self-worth, loved, respected
Isolated, unhappy, unconfident, unfair	Happy, fulfilled
Hurt, dismayed, confused, betrayed	Valued, needed, loved, safe
Unsure of myself, lacking confidence, dependent	Able, creative, healthy, alive

Frightened, can't take anything in	Relaxed, involved, warm, comfortable
Different, stressed, afraid, unloved, ridiculed	Safe, calm, hopeful, light hearted
Insecure, low self-esteem, inadequate, not good enough	Involved, wanted, able to contribute
Embarrassed, anxious, depressed	Valued, strong, confident, human

The affective effects of exclusion are many (Murray 2006, 40).

Disabled people's responses to others' staring can often be at odds with starers' expectations, especially when staring is paired with questioning and other more overt forms of curiosity. There is an "ableist trope of the angry, bitter crip" who refuses to entertain others' curiosity, often embodied through stares (Goodley et al. 2018, 207). This trope exists owing to and in contrast with what Reeve calls the "grateful disabled person role" (2006, 98). It is amplified when the starers understand their own curiosity as morally innocuous. Being subject to the 'angry crip' trope can, in fact, amplify disabled people's anger because it delegitimizes it, suggesting that it is not well-founded anger with a logical and established target, but is instead merely the way *some people* (i.e., they, disabled people) *tend to be*. Although much of what has been said in this section assumes that the disabled person in question — the one who is being stared at — is visibly[9] disabled, similar feelings can arise when someone whose disability is less apparent to others hears ableist claims, particularly when such claims are made by someone who does not know that they are in the company of a disabled person. Moreover, the very act of concealing one's disability can have similar negative emotional outcomes because it assumes the (potential, threatened) presence of a starer. The voice of the starer is internalized, even in the absence of any *actual* starer.

Internalized and Lateral Ableism

The idea of internalized oppression more generally shapes the affective experience of ableism for disabled people too. Internalized oppression is, at the risk of defining it

[9] For more on the distinction between visible disabilities, and on the phenomenon of 'passing' with respect to disability, see Samuels (2003) and Davis (2005).

tautologically, the idea that some oppressed persons come to view themselves in the same ways that their (external) oppressors view them; they *internalize* such viewpoints. For instance, "if women are surrounded by people who view them as subordinate, incapable, or lacking control over their actions," writes Nabina Liebow, they "are likely to come to understand themselves in a similar way" (2016, 715). The same holds for disabled people; if they are surrounded by people who take negative, ableist attitudes toward them (including those attitudes that I highlight in the subsequent section), these attitudes can be taken up reflexively. Ableist attitudes come to *infiltrate* disabled people's consciousnesses (see Nelson 2001). If, for instance, they are surrounded by those who pity them and view their lives as not worth living, these same attitudes — of the ilk Sandra Bartky describes as "messages of inferiority" — can get adopted by disabled people (1990, 23). Internalized attitudes need not be conscious, or they may even coexist consciously with contradictory attitudes; psyches are messy. In her writing on women's psychological oppression, which she defines as "be[ing] weighed down in your mind…[and] to have a harsh dominion exercised over your self-esteem," which accords with definitions of internalized oppression, Bartky writes that women often experience *shame* at their womanhood (1990 22; 30). This shame arises by internalizing the sexist idea that women are inferior to men, or it can be even more removed and be a shame at having internalized the idea that women are inferior to men.

With respect to disability, Campbell and Stramondo (2023) note that "self-stigma" can be a direct fallout from "public stigma," and that self-stigma "has been shown to lead to decrements in hope and self-esteem, loneliness and social withdrawal, and even suicide" (1643). The 'public stigma' that the authors are concerned with has to do with the stigmatizing attitudes engendered in conversations about reproductive technologies, as such conversations often make space for the impulse (or even the obligation) to select against disability. Even outside of this narrow context, the nature of public stigma remains much the same, although the psychological effects on disabled people might be amplified when the decisions being weighed are (literally) life-or-death for people like them. This stigma does not even need to be actualized, though; Nario-Redmond et al find that "the mere anticipation of being stigmatized or socially devalued can increase personal distress," thus showing that the negative emotional consequences of experiencing ableism can persist even absent discrete experiences of ableism (2019, 737).

Reeve identifies two distinct forms of internalized oppression that disabled people face: "false consciousness and double consciousness" (2019, 110). False consciousness seems to align with the more traditional views of internalized oppression, because it pertains to individuals who internalize (that is, come to believe) stereotypes about groups to which they belong. Reeve likens double consciousness to Iris Marion Young's notion of cultural imperialism, a form of oppression[10] whereby the oppressed "experience how the dominant meanings of a society render the particular perspective of one's own group invisible at the same time as they stereotype one's group and mark it out as the Other" (Young 1990, 58-9; Reeve 2019, 109). Rather than fully incorporating the oppressor's viewpoint as in straightforward cases of internalization (false consciousness), double consciousness involves a "*continual*" sense of "presence-as-alien-being-in-the-world" (Reeve 2019, 109). That is, double consciousness might include the false consciousness, but it is paired with a recognition of it as such, because it represents disabled people's coming to be defined both by the dominant culture and by their own "subordinate" one.

Much of what has been said so far has studied internalized ableism as self-directed but externally motivated. That is, I have looked at the ways that disabled people internalize external messages about disability and apply these harmful messages to themselves. There are two further caveats I should make here. First, disabled people can be ableist toward *other* disabled (or presumed-disabled) people. This phenomenon, which has come to be known as *lateral ableism*, muddies the typical power dynamic assumed in instances of ableism. Moreover, in what might seem paradoxical, nondisabled people can be ableist toward other nondisabled

[10] Chapter 2 will provide more detail on how Young uses the term 'oppression.'

people. They might take anti-disability messages and use them to police the behaviors of others who, though not 'technically' disabled, display characteristics or behaviors that are associated with disability (including, for instance, psychiatric disability and neurodivergence). The relevance of lateral harm, or harm among equals, will become more apparent in Chapter 3's discussion of affective injustice, as I believe that other theorists' emphasis on affective injustice following from (other) social injustice might be unnecessary.

This section has detailed *psycho-emotional disablism,* the ways that ableism comes to be *felt* affectively. Ableism can appear psychically, both from overt psychic oppression and internalizing others' attitudes toward disability. Even though psycho-emotional disablism is probably felt most acutely by disabled people themselves, its counterpart — the psycho-emotional dispositions of (those who perpetuate) ableism — cannot be overlooked. In fact, these dispositions might be especially important because of their underpinning more explicit, enacted forms of ableism. I discuss three such dispositions in the next section.

Feeling Badly: Ableist Attitudes

From where does psycho-emotional disablism arise? In this section, I highlight three distinct 'emotions of ableism': pity, fear, and awe, studying the ways that these emotions exist in the minds of (especially) nondisabled people and are applied against disabled people. This is not intended to be an exhaustive list, and it, importantly, includes both two traditionally negative emotions (pity, fear) as well as one that is often considered positive (awe). In what follows, I explore the ways that each of these affects is justified and enacted in the ableist imaginary, and how they express the kinds of sentiments leading to some of the emotions of exclusion outlined by Murray in the table above. What unites these three emotions (pity, fear, and awe) is that those who adopt them in service of ableism often believe their own motivations to be laudable if not at the very least logical. Nevertheless, each of these are instances of what I call 'feeling bad*ly*" — morally condemnable emotions. Ableist attitudes, even when they remain unexpressed, can cause harm because these attitudes shape their bearers' interactions with disabled people in the same ways that implicit biases reinforce oppression. This holds even when such attitudes are expressed exclusively *among* nondisabled people, or when they are unexpressed altogether, rather than only applying when they are made explicit.

Pity

Pity comes to some as a natural response to disability. Pity is often understood as a recognition that others are suffering, and that we (or the pitiers) feel they ought not be. This could be because we feel that this specific form of suffering is undeserved by this specific person, or it could reflect a more general feeling that suffering ought to be avoided. As I show in this section, and as a former disability poster child writes, "pity," no matter how well-intentioned, "oppresses" (qt. Shapiro 1993, 12).

Pity is often wrapped up in a medical model of disability. The medical model of disability maintains that disabilities are best understood as flaws within individual bodies or minds, as "purely mental or physical problem[s] to be cured or treated" (Schalk 2022, 34). Medical models can differ considerably among themselves; they might look at disability in terms of a departure from typical functioning, or in terms of an inability to perform 'typical' tasks (see Barnes 2016, ch. 1 for more on this taxonomy). In short, the main idea driving the medical model is that there is something wrong with a particular body, and the avenues for addressing that wrongness come from (typically allopathic) medicine in the forms of pharmaceuticals, surgeries and other procedures, braces and mobility aids, and so forth. This model is often described as *individual* or *individualistic* too because the targets of its descriptions (and thus of its interventions) are (purportedly) objective facts about individuals.

As Stramondo (2010) suggests, one of the most pressing issues facing disability studies scholars writing against pity is that pity, to the untrained onlooker, seems to be a praiseworthy response. It seems like an achievement of empathy, a signifier of altruism or kindness. Even if we do not totally abandon the idea of pity as "an altruism for the pitier," we must acknowledge that in disability contexts it often remains "an injury to the pitiful" or pitied (Stramondo 2010, 121). The injurious nature of pity, the fact that it causes (or constitutes) harm, should serve as a prohibition against it. Stramondo argues that theories that underscore the value of pity fail to account for the role of power relations in shaping “pitiable harms,” often given as the justification for pity (2010, 124). Either way, he claims, such pity often fails to account for “the contingency of [the pitied party’s] suffering,” making this suffering seem natural instead of created, or it overestimates it, imagining suffering where there is none (ibid). Disabled people are thus understood as perpetual victims, stripped of at least some of their individual subjectivity (Hughes 2019, 94). Misplaced pity like this relies on a medical model of disability because it assumes either that the pain and suffering (where there is some) that disabled people experience is caused directly by their impairments, their physiological conditions and diagnoses. More broadly, it might assume that there *is* such suffering to begin with, despite evidence suggesting that at least some disabled people report quite good quality of life (Albrecht & Devlieger 1999; see Stramondo 2010, 127). Although headlines about the relationship between disability and wellbeing (which seems more or less coextensive with quality of life) might be inconsistent, Barnes notes that “positive disability identity, disability acceptance, and a sense of social acceptance are major predictors of quality of life in the context of many conditions we traditionally consider physical disabilities” (2023, 183). This might not hold for degenerative disabilities, psychiatric conditions, or for those associated with significant ongoing pain (e.g., migraine disorders, osteoarthritis), as Barnes notes, but this still does not mean that sweeping claims about disabled people being worse-off than nondisabled people are warranted (ibid, 186).

Just stating that this pity is ill-informed is not enough to make an argument against its moral value, though. What makes this pity actually harmful against disabled people — what cements its status as *affective weapon* — is that it conveys that its targets have a lower social status than others. Even well-intentioned pity does this, because we don’t see as deserving of pity those who we see as better-off (locally or globally) than ourselves. Pity’s targets are viewed as “damaged goods” (Shapiro 1993, 15). Pity can also be harmful when it contributes (often via its commitment to the medical model) to what Alison Kafer calls “a politics of endless deferral,” the tendency of institutions to get so wrapped up in curing a particular condition that they end up overlooking the more immediate and morally pressing needs of those with it (2013, 29). Pity functions to defer especially when nondisabled people adopt it as an attitude instead of listening to and taking seriously disabled people’s concerns. This reflects a relationship, too, between pity and charity, because charitable giving is often viewed as a relatively low-effort and low-commitment means of helping disabled people — who are always-already framed as in need of such assistance.

Giving can be invaluable, to be sure, but forcing disabled people to rely on others’ kindness forces them to make themselves properly pitiable (Goodley 2014, 121; Hughes 2019, 93). The process of *coming to be seen as an object of pity* can be harmful for disabled people. Consider, again, the relationship between pity and charity. In order to be the recipient of charity, an individual must be piti*able*, and must present themselves as such. Joseph Shapiro documents the evolution of telethons showcasing disabled people, writing that the telethon “remains fixed on the fear of disease, [and] the misery of disability,” because it presents children with “their smiling facings and cheery personalities” which ostensibly “make their impending tragedy all the more pitiable” (1993, 23). That these children are disabled is viewed as an unambiguous tragedy because of cultural scripts around children’s innocence. Disability rights activist Evan Kemp notes: “the handicapped child is appealing and huggable—the adolescent or mature adult is a cripple to be avoided” (Shapiro 1993, 22). The child is thus more pitiable than the adolescent or the adult. The pity approach means that some disabled people are not viewed as pitiable, and others need to make a concerted effort to *make themselves seem pitiable*, itself a taxing endeavor.

Pity is thus weaponized in service of ableism. However well-intentioned, it expresses to disabled people that their ‘condition’ is one that others feel badly about. To be sure, there are times when a disabled person themselves might appreciate others’ pity (see Boleyn-Fitzgerald

2003). The kind of pity I am taking issue with is a sort of blanket pity, a pity that is applied uncritically to many people who are (or are at least perceived to be) disabled without regard for their ability to self-narrate.

Fear

Fear, like pity, serves as an affective weapon within the ableist imaginary. Generally speaking, fear refers to an object-directed[11] anticipatory dislike. That is, to fear is to identify some discrete thing that will come to be — even that will likely come to be — and disavow it, casting it as undesirable. Fear is future-oriented, then. Some fear also has to do with uncertainty; not the assumption that an object *will* come to be, but the possibility that it might. Fear can even be "fictive" in the sense that its object might never come to be, and even believing that it could come to be might be a mark of irrationality (Svendsen 2008, 36). To be sure, there is some slippage between pity and fear. Patrick Boleyn-Fitzgerald identifies *fearful pity* as a particular species of pity that is motivated by a healthcare worker's "fears that his own life will turn out the same" as his patients' (2003, 11). That is, fearful pity is characterized by a *fear* of becoming like the *pitied*, reinforcing the relationship between being pitied and occupying an undesirable social status. Fear, then, is often grounded in some of the same base assumptions as pity. Misguided quality of life assessments, for instance, can amplify the fear of becoming disabled, particularly when coupled with the assumption that being or becoming disabled entails being excluded from public life.

Susan Wendell identifies the "myth of control" as a key driver of the collective belief that disabled bodies should be "feared, ignored, despised, and/or rejected" (1996, 85). The myth of control's target is the body. The myth, then, consists of the belief (perhaps even the aspirational belief, the hope) that the body can be controlled and thus steered away from disability, illness, and death (Wendell 1996, 93-4). The medical institution bolsters this myth through its rhetorics

[11] See Svendsen 2008, chapter 2.

that champion cure and adaptation over accessibility. Among the promises of medicine are the ability to prevent and repair disability, to bring the body back under control (Reynolds 2017, 149). Eli Clare, who writes extensively about the logic of cure that has been forced upon him from birth owing to his cerebral palsy diagnosis, locates beneath the ideology of cure "the fear of body-mind change, aging, and death" (2017, 11). Wendell remarks that even death has been medicalized, leaning into the myth of control. If physicians are prohibited from "record[ing] old age as the cause of [a patient's] death," they might give the impression that death is avoidable, or that its circumstances are at least within patients' and physicians' control (Wendell 1996, 96).

The moralization of disability, too, is bound up in the fearful myth of control. The moral model of disability maintains that it "is a defect caused by moral lapse or sins," religious or otherwise (Olkin 2002, 133). Avoiding disability thus becomes about neutralizing a threat not only to the body, but to the purity of the psyche (Hughes 2019, 92). Succumbing to disability according to this view (on which disability is something to be *succumbed to* in the first place) means ceding control and allowing the 'bad' of the world or of the self to win. If disability is viewed as this sort of moral failing, then disabled people become feared not only for the lack of control their bodies display, but for their character too. This fear also works to make nondisabled people feel more secure in the permanence of their able-bodied status. If people tend to think of themselves as good or at least not-bad-people, and if they buy into some version of the idea that bad things resist happening to good people, *and* if they believe that disability is a bad thing to happen to someone, then they might view themselves as unlikely to become disabled. Their motivation for acting righteously thus may be explicitly tied up in the fear of becoming disabled. This is an extension of Reynolds' "ableist conflation" of "disability with pain and suffering," mentioned earlier in this chapter (2017, 150; 154). Taking a simplified and colloquialized view of pain: if disability inevitably causes pain and suffering and is caused by some sort of self-inflicted suffering, then disability should be feared both for its consequences (pain) and for what it represents about a disabled person's moral character.

Fear, then, is a weapon used against disabled people in that (like pity) it expresses and reifies lower social status. Disabled people are told that they and their bodies are fear-worthy, that there is some innate flaw to them that others are better off for lacking. The entirety of disabled *being* becomes wrapped up in suffering, irrespective of whether this suffering might be outweighed by or counterbalanced with pleasure.

Fear of disability might beget different social treatment than pity. The objects of pity are viewed as unfortunate but are not typically despised or cast out as the objects of fear are. The fear I am concerned with here is not only a fear of disability in the abstract, but a fear of disabled people. Qua objects of fear, disabled people become beings that must be subdued to neutralize the threat they present. Consider the case of Terrance Cottrell Jr., an autistic eight-year-old killed in a "spiritual healing session" — an exorcism — in 2003 (Davey 2003). The comparison here between a disabled person and a demon, an object of fear, is not even metaphorical; the two are literally conflated. Bill Hughes, too, looks toward the material impacts of ableist fear, writing that fear "manifest[s] itself in the segregation of disability, its enclosure behind high walls," referring to the longstanding practice of institutionalizing disabled people (Hughes 2019, 93; see Espinosa 2018). Institutions[12] (asylums, psychiatric hospitals, etc.) were erected ostensibly to protect the people housed in them, but also in large part to protect society from their residents, particularly those taken to be "lunatic, distracted, manic, mad, insane, crazy, mentally ill, psychotic, schizophrenic" (Noll 2018, 307). Those who occupied institutions were viewed as "dangerous because they posed [a] threat to society," and were feared accordingly (ibid, 317). The conditions of institutionalization were horrid, as exposed by lobbyists, activists, and journalists like Dorothea Dix, Elizabeth Packard, Burton Blatt, and Geraldo Rivera (Noll

[12] Noll writes that institutions represent "the confluence of science and religion," of "enlightened scientific thought and evangelical Protestant benevolence" (2018, 309). Thus, pity may be as operative here as fear, because of religious incentives for looking downward upon and trying to improve the lots of the poorly off.

2018). Even generations after Dix's reform activism, institutions remained places in which "physical violence[13] and psychological trauma" abound (Conroy & Downey 2020, 59; 70).

The consequences of ableist fear are thus not limited to the ableist imaginary, nor are they limited to how they make disabled people *feel*; they are material. This fear can result in physical bodily harm to disabled people, and to the 'warehousing' of disabled people in institutions even centuries after activists began warning of the harms such institutions inflict. The harms of institutionalization include the mere fact of segregation and its expressive harms as well as the secondary abuses that happen within institutional spaces. Moreover, because fear often manifests through avoidance, it can be difficult to recognize any particular act as an act of fear. Sometimes, fear might look like the absence of an act owing to the absence of contact. This is one of the reasons I suggest in the following chapter that we need an affective ethic that can account for non-expressed harms.

Awe

Like pity, and perhaps even more-so than pity, awe is often well-intentioned. It is common online and offline to see people marveling at disabled people's ability to adapt, to 'overcome' their disabilities and live full lives 'despite' the hand they have been dealt. In this final subsection, I show that awe, like pity and fear, can be a weapon used against disabled people. I focus on narratives about 'overcoming' disability to show how even ostensibly positive emotions can harm their targets (here, disabled people), thus rendering such emotions — awe, wonder, even a certain sense of pride — potential affective weapons. These narratives might appear to be harmless celebrations of someone's accomplishments, but they rest upon and further convey harmful messages.

[13] In their cataloguing the abuses that occurred at Pennsylvania's Pennhurst State School and Hospital (1908-1987), Conroy and Downey note that institutions' residents were often used for medical experimentation. This shows the compounding fears of disability; those who were feared (because disabled) were subjected to medical experiments indicative of society's fears of *further* illness or disability. In other words, those who were feared were further rendered test subjects toward allaying those same fears.

One of the most common places where awe appears alongside disability is in popular media; film and television, advertisements, memes, and so forth. This is particularly true of the species of awe-provoking media that has come to be known as “inspiration porn,” defined by Stella Young as the pairing of images and videos of disabled people doing fairly mundane things with captions that are meant to “inspire” or “motivate” the audience (2016, 4:10). Such media is often tailored toward a nondisabled audience, making its objectification of disabled subjects even more apparent. Margaret Price writes about the ways that popular audiences might take up a story about disability with (or even *owing to*) an “appetite for stories of ‘tragedy’ and ‘overcoming,’” and Eli Clare describes the ways that “overcoming *bombards* disabled people,” both pointing to the ways that narratives about overcoming are often imposed from externally (Price 2011, 176; Clare 2017, 8, emphasis added). Stories about overcoming make their object — the overcomer — into the object of awe, the person who against all odds defeats their own body/mind. If disabled people are not even given a choice in how they are represented *to* nondisabled people, we can safely assume that these representations are *for* the nondisabled. This can play into one of the elements of weaponization, because the harm/benefit calculus is stacked in favor of the nondisabled.

Price writes that overcoming narratives suggest a sort of brokenness that must be overcome; thus, these awe-inducing narratives are, like feelings of pity and fear, grounded in an individualistic and medicalized view of disability (2011, 52). By overcoming their disability, the disabled person’s broken body is effectively healed. Like the emotions mentioned above, this essentially eradicates disabled futures. If overcoming disability is to be praised, then disability is painted as something without which the world is better. Alison Kafer writes that “curative time” is that ableist temporality which operates on “an understanding of disability that not only *expects* and *assumes* intervention, but also cannot imagine or comprehend anything other than intervention,” including disabled people continuing to exist as such (2013, 27). Centering a future without disability, too, reinforces the politics of deferral mentioned above. By focusing on disabled people's hypothetical overcoming, we might end up erasing their present needs. Resources are deferred because of the focus on a future in which they might not be needed. That these narratives harm disabled people — both psychically and perhaps materially by contributing to the deferral of present needs — renders them weapons.

Awe in this sense can be equated with hope. Hope, like awe, is not harmful *per se,* but becomes harmful when it is applied in certain often patronizing ways. Clare details interviews with Christopher Reeve, an actor who became a quadriplegic owing to a horseback-riding accident. In an interview, Reeve is quoted as saying "There are some people who just really don't dare to hope" (qt. Clare 2017, 10). Hope plays into the same logic as awe, because those who aren't (yet) awe-worthy are expected to hope to become so.

Illustrating the ways that awe-inspiring narratives perpetuate ableist harms, Price cites a 2005 column by Emily Toth in the *Chronicle of Higher Education*. Toth, writing under the pen name 'Ms. Mentor,' "salute[s] *courageous* professors with disabilities who have made themselves as dynamic as the nondisabled," referring to one faculty member in particular "who *heroically*, and secretly, performed kidney dialysis for herself in her office," only publicly disclosing her disability after receiving tenure (Price 2011, 104; Toth 2005, emphasis added). Though Toth does not explicitly invoke 'overcoming,' she does imply it by framing disabled faculty members as energetic *despite* their disabilities, which (we can infer) is supposed to make them less so. Toth does acknowledge that an ideal academy would be one in which disability does not need to be concealed in the first place, but her writing still suggests that (barring this sort of systemic change) concealment is not only a wise option but is a morally praiseworthy one. Of course, such attitudes are rife with weaponizability, because being a *good* disabled person, possessing the right public-facing 'spirit,' is presented as something that all disabled people should aspire to.

These narratives become even more moralized (and thus even more weaponized) through the association of overcoming and concealment with hard work. Kafer and Clare reference a

Foundation for a Better Life[14] billboard, reproduced below, featuring Whoopi Goldberg and the text: "Overcaem dyslexia. HARD WORK."

A billboard depicts an exasperated-looking Whoopi Goldberg with her hands in her hair, looking into the camera, against a purple background. The left side of the billboard reads "Overcaem [sic] dyslexia. HARD WORK. PassItOn.com".

Coy misspelling aside, the pairing of the billboard with *"HARD WORK"* "suggests that other people with dyslexia and learning disabilities who have not met with similar success have simply failed to engage in hard work," and that (thus) being disabled by a learning disability might just indicate a personal failure (Kafer 2013, 8). This too involves assuming that a given disability *can* actually be overcome. Clare likens overcoming to "transcending, disavowing, rising above, conquering" (2017, 9). If we accept that disability is a fundamental part of disabled people's being, how could they possibly be expected to transcend such an essential characteristic? Such expectations are more or less calls to occupy a self that does not exist. Clare was born with cerebral palsy, so the idea of him without it is more or less a fiction. If overcoming is a sort of

[14] PassItOn.com

restoration to a previous (pre-disability) state, such a state does not exist for Clare and others with congenital conditions.

Overcoming narratives harm disabled people in part by measuring them against standards that they (in many cases) cannot meet (see Kafer 2013, 95). This holds true even when overcoming narratives are framed in entirely positive ways, and when a single disabled individual does not take issue with the ways that they are being made the object of nondisabled awe; my concern is with the broader narrative conveyed. What remains flawed about them is the connotations of overcoming as transcendence of disability itself, rather than the disability remaining but some external obstacle being overcome.

What is a Weapon?

Throughout this chapter I've gestured toward the notion of weaponization, that is, of particular emotions being weaponized owing to the role they play in sustaining ableist ideology. We are now at a point where a more systematic account of weaponization is in order. First, a note limiting the scope of weaponization: in order for something to be weaponized, or perhaps even to be weaponizable, it needs to have non-weapon uses. It is not quite that we cannot weaponize weapons, but that to speak of the weaponization of, say, a bomb strikes me as redundant. "This bomb has been weaponized" suggests that something has happened or been done *to* the bomb, when instead its natural state is as a weapon. Bombs are not weaponized; they are weapons.

Second, any act of weaponization should be to the benefit of — or to the intended benefit of, or even merely likely to be to the benefit of — the person doing the weaponizing, and to the detriment of the person against whom the repurposed weapon is wielded. Weapons cause harm, in other words. Intentions and outcomes need not always coalesce. This is not to say that it is always easy to deduce who is benefitting and who is harmed by a single act, but that, generally speaking, weaponization serves to protect or benefit someone while harming another. This is true of the mere weapons mentioned in the previous paragraph, too; they straightforwardly are only supposed to be weapons *against* one party. They are weapons *for* (the benefit of) the other party. This is captured by the contention that emotion weaponizing equates to using emotions to disadvantage people (Pismenny et al. 2024, 166).[15]

More substantively, I take this second requirement to be a requirement of *harm*. This harm can be intentional or unintentional, conscious or un/sub-conscious. Harm, importantly, can also be instantiated between or among equals. There is no requirement of disparate social standings between parties for an action to count as harmful. Following Molly Gardner (2021), I view the currency of harm as well-being; that is, harm results in a net decrease in well-being, and so whether an action is harmful depends on whether it causes or is intended to cause a decrease in well-being, however that might be defined. Without getting into the specifics of different accounts of well-being, to say that harm inflicts damage to one's well-being is to say that it makes its victims "comparatively worse off than before," or that it renders them "badly off" in a non-comparative[16] sense (Unruh 2023, 896; Dunne & Kotsonis 2023, 3; see also Purves 2019). This harm might consist of 'merely' being made to feel bad (sad, angry, or otherwise distressed) or it might be more material.

Third and finally, weaponization need not be apparent to either (or any) of the relevant parties. In *Dogwhistles and Figleaves*, Jennifer Saul identifies two types of dogwhistles: Overt Code and Covert Effect. Overt Code dogwhistles "work like a secret code," used by the group of those 'in-the-know' about the dogwhistle (Saul 2024, 9). Covert Effect dogwhistles, on the other hand, "act on some of the people who accept the anti-racism norm without their awareness of this," that is, without their being aware that the dogwhistle is working on them because of their deep-rooted racial resentments (ibid). The key difference between the two kinds of dogwhistles, then, is whether the group for whom it is a dogwhistle actively or explicitly endorses the (e.g.) racist attitude that the dogwhistle stokes. If we think of dogwhistles as a kind of weaponized

[15] More will be said in the next chapter about Pismenny et al.'s account of emotional injustice.
[16] Because she allows for both comparative and non-comparative detriments to well-being, Unruh describes her own account as a "hybrid" one.

language, then it is apparent that the hearer of a dogwhistle — which may include those targeted by it — need not be aware of the language's weaponization. Moreover, if they repeat the language, they can inadvertently become the dogwhistle's user; the *weaponizer* (Saul 2018, 368).

These desiderata might be troubling because collectively they — the requirements that weapons be weapon*ized*, that they cause harm, and that they need not be apparent to either party — might leave open to definition as weaponization cases that we might want to grant as morally neutral or even positive. For example, if someone is acting in an abusive manner, but their target knows that this person will be responsive to emotional appeals, the targeted person (the abused) might intentionally invoke pity to try and get the abusive person to reconsider their words or actions. This case, in meeting the criteria above, might indeed be classified as weaponization. Unlike many of the cases I consider throughout this dissertation, though, I do not see it as *wrongful*. That is, it lacks the normative weight that leads us to say that other cases are morally impermissible. Thus, we cannot paint in too broad strokes about weaponization; it need not always be wrongful. The cases I consider above are cases of weaponization insofar as they result in an individual or group of individuals feeling *bad*.

Conclusion

Harm is not limited to physical violence. Emotions, however well-intended, can have psychologically and materially harmful effects, even when they remain unexpressed. This is especially evident in the case of disability sentiment, because disabled people are often made into the subjects of cultural lore non-consensually. That is, even without personally interacting with a single disabled person, someone can come to project narratives onto the entire disabled community. These narratives can be negative — the idea that disability (and thus the disabled subject) should be feared — or they can consist of more beneficence-tinged stories about pitiful tragedies that are overcome through sheer power of will. The affective dimensions of ableism, though, are not limited to the feelings or responses that the nondisabled have about and to disability. They include the affective fallout that disabled people are left to grapple with, as well as the ways that the mere threat of disability — or of being treated as disabled — manifests between individuals with the same disability status. In this sense, ableism itself is a weapon insofar as it wears on disabled people *and* insofar as it creates ableist feelings (*bad feelings*) even absent the one-directional power structure assumed in conversations about ableism. Ableism weathers and, in turn, subjects its objects to even further harm.

The idea of emotions being a currency of harm, of ethical transgression, is not novel. Examples abound in literature (academic and otherwise) about anger being used to harm, and about the ways that being treated poorly psychologically harms victims both with and without the background of some grander social injustice. In the subsequent chapter, I home in on the nascent yet growing literature on *affective injustice*. Affective injustice attempts to systematically make sense of the relationship between social injustice and the emotions and, moreover, describes the ways that the emotions can be either the target of injustice or its primary tool. Some of its theorists even explicitly name emotional weaponization as an injustice. Chapter 2 serves as a comprehensive overview of the affective injustice literature, as well as shortcomings that I identify within it. In the following chapter, I note that the affective injustice literature as it stands might not be able to account for injustices like those described above, particularly when the injustice can (in theory and in practice) remain unspoken and (thus) without uptake.

Chapter 2 - Affective Injustice and its Discontents

Introduction

As established in the previous chapter, there are compelling reasons construct an affective ethics. Emotions and harm are undoubtedly related; anyone who has been harmed by another (so, *anyone*) can attest to having experienced the emotional or affective consequences of their mistreatment. Mistreatment hurts, it feels bad, but ethicists should task ourselves with discerning what actually makes a given instance of feeling bad wrong; when does it rise to the level that an affective ethics should be properly concerned[17] with? That is, the fact that something leads to negative feelings does not necessarily entail that it is harm, just as something's being harmful does not necessarily tell us that it is *unjust*. How, then, are we to craft an affective ethics without relying on (misleading) claims about negative affect alone being sufficient for identifying transgressions? Moreover, how can we talk about 'unethical' affective behavior without categorizing all such behavior as unjust?

In recent years, the affective injustice literature has come about as an attempt to outline criteria for when our encounters with others' emotions are unjust, rather than merely unpleasant, which is considerably less normatively loaded and seems to beget weaker calls for intervention. Those who offer definitions of affective injustice identify a set of guidelines for judging when one party's treatment of another's emotions is unjust. This chapter surveys the affective injustice literature, including a brief look at the traditions that it builds upon and stands in relation to. Much of this work is translatable to conversations about affective ethics (rather than [in]justice) more broadly. That is, while the cases that many of the authors below

[17] For an analogy: just as someone's being sick does not mean that a violation of medical ethics has occurred, someone's experiencing a negative emotion does not mean that a violation of affective ethics has necessarily occurred.

outline are enacted against a background of (or because of) social injustice, they create space for a broader consideration of the ways that individuals can act unethically with respect to others' emotions and their own emotions. This includes, for instance, mistreatment among equals, where neither party is situated in a way that allows them to oppress the other. As noted in the introduction to this dissertation, I share Sarah Schulman's hesitancy to label all conflicts as abuses, or all acts of mistreatment as injustices. Nonetheless, much of what follows uses the language of (affective) injustice in keeping with other authors' terminology. I critique this loose use of 'injustice' in the penultimate section of the chapter, suggesting that the insights about affective injustice apply as well to affective wrongdoings.

I begin this chapter by providing a genealogical account of affective injustice, exploring how it developed out of other paradigms including the rich history of feminist writings on psychological oppression (itself an attempt to concretize the relationship between structural injustice and emotion), as well as the more contemporary literature on epistemic injustice, which grounds (at least methodologically) much of the affective injustice literature. I also consider affective injustice's relationship to more diffuse ideas like emotional injustice. Next, I provide an overview of extant accounts of affective injustice. The affective injustice literature is still relatively sparse, given that the term only came into use in 2018 in separate papers by Amia Srinivasan and Shiloh Whitney, but the handful of papers that use it do so in ways that have important differences. In other words, even if we grant that there is a singular affective injustice *literature*, this literature consists of authors who conceptualize their target of inquiry quite differently. Third and finally, I measure the efficacy of the extant work on affective injustice by judging its ability to capture the phenomena described in the previous chapter, as well as its ability to capture the wrong of lateral harms that might not rise to the threshold of injustice. My contention is that the extant accounts are uptake-based (or at least based on *hypothetical* uptake), and so cannot take account of ableist attitudes that might not actually be expressed and (therefore) not taken up. Moreover, I am not sure that current theories fully explain why the cases they identify should actually be considered *injustices*, a considerably higher hurdle to clear than, say, *affective harms*, given that they can theoretically occur between similarly situated people. This section serves as a transition to chapter 3, wherein I am concerned with other discrete cases of non-uptake-based affective injustices, namely, the promulgation of social norms that dictate acceptable responses to others' emotional claims (especially given that such responses can exist irrespective of a particular emotion's being given uptake).

Case Studies in Affective Injustice? Disabled Rage(s)

As I will be providing in the subsequent sections an overview of the affective injustice literature as well as its shortcomings, it can be helpful to have a specific case in mind. In the next chapter, I show that extant accounts of affective injustice cannot adequately account for the weaponization of therapy-speak, but I will use different case studies throughout this chapter so as not to beg the next chapter's question. In the interest of continuing the previous chapter's discussion about affect and ableism, the cases I will use here are ones that I call *Disabled Rage* (DR). I will address *Disabled Rage* through two separate cases: DR1 and DR2. DR1 is a case study that I find the main accounts of affective injustice (Srinivasan's, Whitney', and Gallegos') largely able to accommodate, though I show how some minor modifications (including the modification of DR1 into DR2) to the case may change this. DR1 is discussed, though not by name, in some of the previous chapter's affective account of ableism. It refers to disabled people's own rage at ableism, rather than encompassing nondisabled people's rage at disabled people (or, for that matter, at ableism).

I model *DR1* after two main theoretics: Sara Ahmed's feminist killjoy, and Myisha Cherry's Lordean rage. Ahmed's killjoy is one who *kills joy* insofar as she refuses to capitulate to violent and oppressive norms of/and culture. "The killjoy," she writes, "is one who does not make the happiness of others her cause," instead choosing to shed light upon "the denial of inequality under the assumption of equality [as] a technique of power" (Ahmed 2017, 75; 252). The killjoy becomes what Ahmed terms an "affect alien"[18] by refusing to experience the 'right' emotions, that is, those of happiness even in the face of oppression (ibid, 57). Similarly, the subject of *DR* is one who refuses to burst with gratitude for basic accommodations and thus seems (at least in the eyes of an ableist social body) to feel wrongly by not feeling grateful enough. Cherry's Lordean rage is grounded in the work of Audre Lorde, particularly her essay "The Uses of Anger: Women Responding to Racism," and targets "those who are complicit in and perpetrators of racism and racial injustice," thus acting as a call for solidarity in the face of such injustice (2021, 23). Lordean rage is pragmatic and action-driven, rather than being (for example) solely destructive toward its bearer or solely focused on retribution.

We can thus understand *DR1* as follows: it is the rage that disabled people experience in coming up against or being forced to come up against ableist structures, including both social and cognitive structures and physical infrastructure. This rage can be expressed toward particular individuals, especially when those individuals are acting in service of ableism, but it can also be a less targeted rage against the social structures that keep ableism and other forms of oppression in place. It is an anger that is directed not into a void, but toward rectifying injustice.

DR2 is different. Consider the following: two political organizers, Jay and Kay, belong to the same organization. They are similarly situated — they are both middle-income white cisgender women — and both have been diagnosed with ADHD. Their organization holds hybrid meetings monthly, which both Jay and Kay enjoy attending in-person, though some members participate remotely and are projected onto a large screen in the front of a room. Jay's ADHD leads her to ask for closed captioning on the projector screen; she has difficulties with auditory processing and finds that the captions allow her to follow what the remote speakers are saying. Kay, on the other hand, finds the captions distracting owing to her ADHD. She cannot merely look away from the screen, though, because important information and graphics are often visually

[18] This notion is addressed in more depth in the following chapter.

displayed on-screen. Just as Jay cannot process what the remote attendees are saying without captions, Kay cannot process what they are saying *with* the captions. None of the other members of the organization who attend in-person prefer or against the closed captioning. Jay begins to resent Kay, to feel immense anger toward her – and perhaps vice versa. This resentment is the crux of DR2.

Unlike DR1, DR2 cannot be attributed to ableism (at least not cleanly[19]) in the sense of a nondisabled person oppressing (or otherwise acting poorly toward) a disabled person. DR2 is not a case of one party's disregard for the other's disability, but of competing access needs. A similar case might arise, for instance, if the group consisted of one person with low vision who needed bright overhead lights and another with a migraine disorder triggered by bright lighting. Though workarounds might not be completely impossible, they nonetheless would require that one party do something they do not want to do or place themselves at a disadvantage. Because neither Jay nor Kay wants to be put in this position, their anger toward one another builds.

Shifting Registers

'Affective injustice' did not arise *ex nihilo*, nor were the phenomena it targets completely alien to feminist philosophers prior to the term's being coined in 2018. In this section, I trace the lineage of affective injustice, drawing connections between it and three related and ongoing traditions in feminist philosophy: epistemic injustice, emotional injustice, and psychological oppression.

From Epistemic to Affective

Affective injustice, as a named concept, can be traced back to papers by Amia Srinivasan and Shiloh Whitney. Though Srinivasan and Whitney theorize affective injustice quite differently, as will be explained further in the following section, both make a point of explicitly noting the

[19] I do not want to separate DR2 from ableism altogether. To do so minimizes the impact of structural injustices on interpersonal relationships. However, I still find it oversimplifying and misleading to suggest that cases like DR2 are cases of ableism simpliciter.

similarities and differences between their proposals and Miranda Fricker's earlier notion of *epistemic* injustice.

Fricker coined epistemic injustice to refer to the myriad ways in which individuals can be harmed in their capacities as knowers. Epistemic injustice consists of two distinct yet overlapping phenomena: testimonial and hermeneutical injustice. Testimonial injustice is the focus of much of Fricker's book and consists of cases where "a hearer give[s] a deflated level of credibility to a speaker's word," owing primarily or exclusively to some "*identity-prejudicial credibility deficit"* (2007, 1; 28, emphasis in original). In other words, testimonial injustice occurs when someone's membership in a social group (on the basis of race, sex, gender, ability, etc.) is taken as justification for finding them non-credible, or non-believable. Simply believing someone to be a liar does not mean that you are committing an injustice against them, because their being a liar is a genuine reason for disbelief. One of Fricker's central cases of testimonial injustice is Tom Robinson of Harper Lee's *To Kill a Mockingbird*. In the novel, Robinson is a Black disabled man falsely accused of raping a white woman. Fricker argues that when an all-white jury fails to believe Robinson's testimony, he experiences a unique type of injustice that targets him in his capacity as a knower by stripping him of credibility owing to anti-Black prejudice. This is testimonial injustice.

Hermeneutical injustice, on the other hand, occurs "when a gap in collective interpretive [that is, hermeneutic] resources puts someone at an unfair disadvantage when it comes to making sense of their own social experiences" (ibid, 1). Hermeneutical (interpretive) resources are, for Fricker, "our shared tools of social interpretation," and consist primarily of language; the words and phrases that we use to make an experience intelligible to ourselves and others (ibid, 6). "Structural identity prejudice in the collective hermeneutic resource" often leads to some persons' experiences not being captured in that resource (ibid, 155). Examples from Fricker's work include the experiences of sexual harassment and postpartum depression prior to the terms entering the public lexicon.

In the years since Fricker's book was published, a vast literature on epistemic injustice has arisen, cementing it as a cornerstone of feminist philosophy. For instance, Gaile Pohlhaus Jr. posits a "truncated or circumscribed subjectivity" in response to Fricker's claim that the result of testimonial injustice is the victim's being "deprive[d]…of a certain fundamental sort of respect," thereby rendering them an object rather than subject of knowledge inquiry (Pohlhaus 2014, 103; 105; Fricker 2007, 132). Pohlhaus' argument is that known objects actually make a claim on their knowers that is denied to victims of testimonial injustice, at least in its most extreme forms. Similarly, theorists like Audrey Yap and Emmalon Davis propose that Fricker's characterization of testimonial injustice as being about credibility deficits alone might be overly simplistic, because there are cases where credibility excesses can be similarly unjust (Yap 2017; Davis 2016). Moreover, some suggest that Fricker's hermeneutical injustice paradigm might not fully account for the social mechanisms by which certain tools of interpretation become marginalized, arguing that Fricker's model relies too much on a simplistic model of exclusion from a single collective hermeneutical resource (Pohlhaus 2012; Dotson 2012; Falbo 2022).

Both Srinivasan and Whitney credit Fricker's paradigm with inspiring their own. Srinivasan writes in a footnote that "the notion of affective injustice bears many similarities to Miranda Fricker's (2007) notion of *epistemic injustice*," though she does not elaborate much further (2018, 136). Whitney's mention of Fricker is also quite brief, though she expands upon the connection a bit more than Srinivasan does, writing that:

> [Marilyn] Frye's description [of women's anger being denied proper uptake] suggests an affective injustice parallel to what Miranda Fricker has termed 'epistemic injustice' (2007). While *epistemic* injustice damages the credibility given one's claims, *affective* injustice damages the weight afforded one's feelings. The weight at issue is not that of belief, but of affective force (2018, 495).

Perhaps something like this is what Srinivasan intended by her mention of the 'similarities' between Fricker's view and her own; they might be structurally similar, but where Fricker is concerned with knowledge (and the requisite notions like belief and credibility), Srinivasan, like

Whitney, is concerned with *affect* or *feeling*. The idea of 'affective force' — what Whitney describes as the affective injustice counterpart to belief — will be fleshed out more fully in the subsequent sections.

What epistemic and affective injustice have in common, though, is that they represent a turn toward the inner dimensions of injustice. This is not necessarily in a phenomenological sense, but more so in the sense that they look at the internal, psychic dimensions of injustice, rather than taking materiality as their starting point. I am, moreover, optimistic that the affective injustice can be taken up as widely as epistemically injustice has. Though it points toward a different aspect of our experiences of injustice, it points to something that seems as omnipresent as knowledge-centered injustices are.

Affective Injustice and Emotional Injustice

Affective injustice — and, as a consequence, affective *justice* — is not entirely like what journalist Esther A. Armah (2022) describes as *emotional justice*, which might nominally suggest the same meaning. Armah understands emotional justice as work that "explores the impact of history's systems of oppression when it comes to the emotional" and tries to rectify the wrongful emotional impacts of oppression, particularly "regarding race, trauma, whiteness, and history" (2022, 4). Her book's subtitle, too — *A Roadmap for Racial Healing* — points to the specific context(s) in which she intended the concept to be deployed. Throughout, she calls for people of all races and across continents to do the necessary "emotional work" to disrupt the legacies of racial injustice (ibid, 31; 37). Armah's work is thus narrower in scope than the affective injustice literature is, to say nothing of its relationship with an affective ethics. Armah's book no doubt makes important contributions to the conversation about the relationship between injustice and emotions, though. For instance, she melds the personal and the political, describing the ways that even intimate relationships are shaped by structural (racial) ideologies (ibid, 51). Some of the examples she considers are similar to those that come up in discussions of affective injustice; for instance, her discussion of "emotional mammies" mirrors Whitney's description (see below) of Black women's being forced to metabolize not only their own anger, but white people's too (ibid, 76).

Moreover, affective injustice is not quite equivalent to the broader phenomenon Pismenny and colleagues refer to as emotional injustice; nor, for that matter, is these authors' categorization equivalent to Armah's. Because she focuses only on *racial* emotional (in)justice, Armah's emotional justice is narrower (though not in a pejorative way) than that of Pismenny et al. Pismenny et al. have a few reasons for proposing an alternative to affective injustice, rather than subsuming what they are writing about under it. Chiefly, they write that "the phenomena under examination" in their paper "generally involve specific discrete emotions rather than mere negative and positive valence, which the term 'affective' is sometimes taken to imply" (Pismenny et al. 2018, 153). That is, they suggest that many of those who write about affective injustice are actually writing about emotional injustices, albeit using the language of affect theory. It is worth noting that when I speak of my desire for an affective ethics, I am using 'affective' in the same way that Pismenny and colleagues suggest the theorists of affective injustice are using it, that is, as more or less synonymous with 'emotional.'

On the Pismenny account, emotional injustice is understood as occurring "when the treatment of emotions is unjust, *or emotions are used to treat people unjustly*" (ibid, 154; emphasis added). This is closer to the target of my intervention than affective injustice (as it is used in the literature) is, as it accommodates a wider range of phenomena than simply injustice done unto (or resulting in negative consequences for) the emotions. Thus, affective injustice as understood by those who use the term explicitly includes some of the phenomena of emotional injustice (outlined by Pismenny et al.) *in addition to* those that Pismenny et al. acknowledge as part of the proper domain of affective injustice.

Affective Injustice and Psychological Oppression

Moreover, affective injustice theorists are not simply recapitulating what feminist philosophers have been discussing for decades through the language of (among other terms) ‘psychological oppression’. Before highlighting the differences between the two — both conceptually and in how they are used in the literature — a summary of the psychological oppression literature within feminist philosophy is in order. To begin, as established in the previous chapter, the idea that harm (including psychic harm) and emotion are related is far from new, particularly within feminist philosophy traditions. In one of the more complete accounts of psychological oppression, Sandra Bartky writes that “to be psychologically oppressed is to be weighed down in your mind,” or to have others exercising undue influence over your psychic states (e.g., your self-esteem) (1990, 22). Psychological oppression can be a sort of self-fulfilling prophecy, too, because the psychologically oppressed come to replicate these patterns of oppression over their own minds. Drawing from Frantz Fanon’s work on psychic alienation, Bartky posits that psychological oppression consists of three categories: “stereotyping, cultural domination, and sexual objectification” (ibid, 23; Fanon 1982, ch. 2). Bartky explains each of these axes (respectively) as follows:

> 1. Women are subjected to stereotypes which “threatens our self-determination” via the shadow of “an alternative self, a truncated and inferior self” (ibid, 24).
>
> 2. Women experience cultural domination in so far as “all the items in the general life of our people…are sexist,” resulting in alienation for those who fail to see themselves in this dominant culture (ibid, 25-6).
>
> 3. Women are sexually objectified, meaning that their “sexual parts or sexual functions are separated out from the rest of the[their] personality…or else regarded as if they were capable of representing [them]” (ibid, 26).

All three of these examples represent a loss of respect for one’s personhood, often enacted through the psychologically oppressed’s (sometimes newfound) belief that they are *genuinely incapable* of fulfilling the demands of personhood (e.g., autonomy, understood as the ability to choose one’s own means of self-actualization).

Another substantial account of psychological oppression comes from Ann Cudd’s *Analyzing Oppression*. Cudd writes that we need a psychological account of oppression in order to make sense of the ways that “the oppressed come to be psychologically subdued to acquiesce or even participate in their own oppression," that is, the ways (and reasons) that oppression appears to be so resilient even in the face of resistance (2006, 55). She focuses, in other words, on the ways that oppression impacts the psyche of the oppressed, thereby rendering them vulnerable to further oppression. She considers psychoanalytic, Hegelian recognition-based, and social-cognitive approaches[20] to this question, as well as responses to such approaches. Landing on an explanation of how oppression becomes codified in the minds of the oppressed, Cudd suggests, helps to make sense of how "the psychological forces of oppression are created by and in turn reinforce the psychology of oppression" (ibid, 80).

Replying to critics who might say that the affective injustice literature is redundant because of what has already been written about psychological oppression, Katie Stockdale argues that (conceptually speaking) affective injustice and psychological oppression serve different ends, and that there is thus reason to support the newer literature on the former. Stockdale writes that "whereas psychological oppression is found in the hearts and minds of people who are oppressed, affective injustice is most fruitfully understood as a *structural* phenomenon" (2023, 2). In this respect, Stockdale seems to be separating out the more comprehensive accounts of psychological oppression (Bartky's and Cudd's, for instance) from the smaller-scale examinations of how emotions are used toward oppressive ends. Moreover, Stockdale is pointing out that affective injustice is a less tangible phenomenon.

Stockdale's characterization should not be taken to suggest that psychological oppression is only or merely about internalization; of course, psychological oppression can result in quite overt outward displays of emotion, as in the "explosive anger" that María Lugones describes (2003, 103). Moreover, the term has been used to refer to phenomena that are not primarily about an individual's being psychically harmed in a way that reinforces their oppression. For this reason, Stockdale suggests classifying internalized oppression "as a specific *form* of

---

[20] Cudd ultimately finds that the "cognitive process of stereotyping" makes possible the sort of social grouping that gets reinforced by oppression (2006, 81).

psychological oppression," albeit not one that represents the entire category (2023, 5). Psychological oppression, she writes, refers to "the ways that oppression takes up residence in people's minds through mental states and processes that reinforce their oppression" (ibid). So, while psychological oppression certainly includes phenomena like internalized oppression (see Jaggar 1989; Liebow 2016), it also includes ones like intimidation (insofar as intimidation represents an oppressor's taking up residence in the oppressed's consciousness, even if the oppressed does not come to mimic the oppressor's thought patterns).

Stockdale suggests a redefinition of affective injustice as "affective norms, practices, and relationships that are embedded in *social conditions* of injustice" as a way of negotiating between the three dominant views of the concept (Srinivasan's, Whitney's, and Gallegos') (2023, 10, emphasis added). The focus on social conditions is what, for Stockdale, differentiates affective injustice from psychological oppression. This seems similar to Fricker's requirement that in order to be properly classified as epistemically injustice, a given (e.g.) credibility deficit must stem from identity-prejudice (2007). So, in order to discern whether a disabled person's circumstances warrant the label of affective injustice, what would matter is not merely whether they are experiencing something that fits the criteria that different authors offer for affective injustice, but whether this experience might reasonably be traced backward to some systemic root cause (in this case, ableism).

Extant Accounts of Affective Injustice

In this section, I provide an overview of the extant affective injustice literature. Though the phrase 'affective injustice' has only been used widely since 2018, a number of accounts of and commentaries upon the topic exist. I begin by exploring the most fleshed-out accounts of affective injustice — those of Amia Srinivasan and Shiloh Whitney — before turning to those that are less systematic but still contribute to the ongoing theorization of this phenomenon. This section is primarily descriptive in function; in the following section I raise critiques of the existing accounts of affective injustice.

Srinivasan

Srinivasan's account of affective injustice frames it in terms of a normative conflict that the already- (that is, otherwise-) oppressed — those who otherwise have grave injustice such as racism, sexism, ableism, and so forth done unto them — experience. Affective injustice is, then, a "second-order injustice that is parasitic on first-order injustice, a sort of psychic tax that is often levied on victims of [other] oppression" (Srinivasan 2018, 135). This adds to the moral imperative of naming and addressing such injustice, since it targets those who are already (and, we must assume, wrongfully) experiencing oppression. If there is an un- or under-recognized form of injustice secondary to those that we are more likely to recognize, then "things are even worse than we generally take them to be" (ibid, 136).

To better understand Srinivasan's account of affective injustice, it helps to start from the example she begins with: a 1965 debate between author-activist James Baldwin and William F. Buckley Jr., "the most important public intellectuals, respectively, in the American civil rights movement and the American conservative movement" of their time, on the relationship between anti-Black violence and the so-called American Dream (ibid, 123). In the debate, Baldwin expresses that Black Americans have great reason to be angry, and while Buckley does not disagree *per se*, he warns that "in the end, Negro anger would be met…with white violence" (ibid, 124). Black Americans must choose, then — if this can properly be described as a choice rather than as a facade of a choice — between bottling up their anger, no matter how righteous it may be, and expressing that anger and risking white retaliation.

This tension between anger's righteousness, or *aptness*, and its (alleged) counterproductivity — for the angry person themselves or for others — makes up the core of affective injustice for Srinivasan. Anger is apt (righteous, fitting) when it is a response to "a *moral violation*:[21] not just

[21] Srinivasan frames affective injustice in terms of its relationship to social injustice. Accordingly, readers might assume that by 'moral violation' she means only injustice. However, I maintain that even more mundane moral violations - like those that arise among equals in DR2 - can be violations in this sense.

a violation of how one *wishes* things were but a violation of how things *ought to be*"; apt anger is not a matter of individual preference-frustration, but of a more universalized or normatively significant sort of preference-frustration (ibid, 128). In addition to being responsive to a genuine moral wrong, apt anger must be responsive to that wrong (and not merely coincidental) and proportionate to the relevant wrong.

Srinivasan thus defines affective injustice as "the injustice of having to negotiate between one's apt emotional response to the injustice of one's situation and one's desire to better one's situation," which is wrongful because "it forces people, through no fault of their own, into profoundly difficult normative conflicts" (ibid, 135; 136). The normative conflict arises when an oppressed group's (apt) anger is used against them. Consider, for instance, the "Angry Black Woman" (ABW) stereotype[22] that Myisha Cherry explores in her response to Srinivasan's paper; typecasting aptly angry Black women as such is harmful not only because it is an unfair evaluation that comes with negative consequences that other groups' anger lacks, but also because it is then taken as reason for Black women to avoid expressing any anger, no matter how apt (Cherry 2023, 136). This holds with both DR1 and DR2 too; disabled people who express anger about ableist circumstances and about the misfit their disability creates between themselves and the world are often accused — mostly by nondisabled people but also by other disabled people — of being ungrateful, nagging, and so on, and so any further disappointment they show is chalked up to their being bitter. These normative conflicts can thus be discussion-enders that foreclose the possibility of further conversation.

The ABW stereotype (an instance of what Cherry terms an "affective stereotype") functions to "bully and blame Black women, as well as silence them and their claims against a patriarchal, white supremacist world" (ibid, 138). The normative conflict here, in the sense that Srinivasan uses the phrase, is between Black women's "getting aptly angry or disconfirming the ABW

[22] Myisha Cherry defines stereotypes as "widely held cognitive generalizations about specific groups" (2023, 136).

stereotype" (ibid, 143). Confirming the stereotype, Cherry writes, carries the internal and extrinsic[23] risks of stereotype threat. A similar thread is echoed by María Lugones, who concedes that "anger needs to be trained *but not necessarily toned down*" (2003, 105, emphasis added). A blanket prescription for toning anger down, like a recommendation that those experiencing either iteration of *disabled rage* internalize their anger so as not to expose themselves to further ableism, would be a textbook counterproductively claim whereas calling for anger to be trained (and to be trained in methods other than suppression) recognizes the transformative potential of anger, including DR1 (see also Cherry 2021). Even barring this potential, that is, even if anger had no power to transform unjust circumstances, it could still be apt or called-for.

To be sure, Srinivasan's account of affective injustice is not limited to contexts of anger. She writes in a footnote that, in addition to experiencing affective injustice when one's drive toward anger conflicts with the desire to better one's circumstances, "one could also suffer affective injustice by being presented with an occasion for apt but counterproductive sadness, hopelessness, despair, etc." (2018, 135 fn. 45). Affective injustice thus lies in the normative conflict rather than in any particular emotion.

Whitney

Whitney's account of affective injustice is rooted in the phenomenological tradition and draws heavily from Merleau-Ponty's work. She begins with Marilyn Frye's claim that women's anger is "generally not well-received," especially by men, and that this ill-reception leads angry women to be characterized as "crazy bitch[es]" rather than as persons making reasonable claims on respect (Frye 1983, 84; 89). That is, women who are dubbed 'crazy bitches' are not always — or often, perhaps — acting unreasonably; they're merely demanding to be respected. Whitney suggests that Frye's claim represents a distinctly *affective* injustice, where rather than a person's

[23] By 'extrinsic harms,' I mean those harms that exist in addition to the persistence of the stereotypical belief (which can be held by the stereotyper or the stereotyped). For instance, consider Monique Morris' writing on how stereotypes of Black girls as sexually promiscuous contributes to their sexual exploitation, which in turn impacts educational attainment (2016, ch. 3).

knowledge claims being denied credibility (as in Fricker's epistemic injustice), "the weight afforded [their] feelings" is damaged because their feelings are stripped of their "affective force" (or, it seems, granted a distorted sort of affective force) for others (Whitney 2018, 495).

Whitney's rendering of affective injustice relies on Merleau-Ponty's "body schema," according to which the boundaries between the internal (bodily) and external worlds become blurred (ibid, 491). The understanding of affect enmeshed in Merleau-Ponty's body schema is one according to which "affects are not 'mere' affects, presenting us only with internal states," but are instead "materialize[d] in a shared world as embodied behavior that radiates affective force" (ibid, 493). By affective force, Whitney seems to mean something like a sort of affective continuity, or the ability for a given affect to both (1) be intentional, or motivated, for the person who is feeling it but also (2) be moving to others; be able to *affect* them.

There are different levels or layers to affective injustice as Whitney understands it. (2) above might be curtailed altogether; if an interlocutor simply fails to respond, or fails to respond appropriately, to my emotions, they have effectively disrupted those emotions' affective force. This is what Whitney refers to as "affective marginalization," an affective counterpart to Iris Marion Young's definition of marginalization as the creation (and sustenance) of an "underclass…a whole category of people [who are] expelled from useful participation in social life and thus potentially subjected to severe material deprivation and even extermination" (Whitney 2018, 495; Young 2011, 53). The relevant aspect of social life that affective marginalization bars its victims from seems to be the often unregistered (because of how taken-for-granted, how unremarkable it is for people in positions of relative power) experience of one's emotions receiving uptake.[24]

In addition to a failure to respond at all, affective injustice as Whitney understands it might look like a person's emotion being reduced to "mere affect," or affects that do not ask or require

[24] It is also noteworthy that here that Whitney seems to take a disjunctivist view of uptake and action; she tries that "giving uptake doesn't require you to capitulate to my anger's demands" (2024, 42).

anything of the affected person's interlocutor(s). Thus, even if an emotion is "validated" on the surface-level, and thus given some sort of cursory uptake, there may still be affective injustice at play if a hearer fails, as Whitney puts it in a more recent paper on affective injustice, to "cooperate with [a speaker's] anger enough for to fulfill its other-orienting function" (Whitney 2024, 38; 42). Uptake, generally speaking, is about sensitization; it is about being moved. In the DR cases, Whitney's idea of affective injustice might look like the subject's anger being ignored (that is, failing to move others at all), or like their anger being recognized but not having any further effects.

Whitney is careful to note, though, that affective marginalization does not capture the full range of affective injustices. We must also, she writes, identify affective exploitation and affective violence, because one's emotions receiving uptake is not enough to guarantee justice; there are several ways that emotions can be taken up (that is, not denied uptake or marginalized), but not taken up *justly*. These other kinds of affective injustice have to do with "the exploitative and violent displacement of affective force," particularly when affective force is displaced onto oppressed persons (Whitney 2018, 497).

To gain a clearer sense of affective exploitation, consider a metaphor that Whitney uses in describing Audre Lorde's writing on anger: the body as "affective waste disposal site" (ibid, 499). Here, she suggests that Lorde (and other Black women, and oppressed persons more generally), not only have to contend with their own emotions, but also with the emotions of others with more power. When Lorde writes about having to bear the brunt of anger, she is not only concerned with her own (apt, to use Srinivasan's verbiage) anger, but with the anger that is a byproduct of racist and misogynistic hatred; she (Lorde) thus has to spend her own affective resources on responding to others' emotions, while those others (in the case Whitney cites, white women in academia) fail to maintain the same regard for her emotions. Whitney uses the phrase “exploitative affective labor” in ways similar to how phrases like “emotional labor”[25] are used in popular conversation, though with a wider scope (ibid, 501).

Lastly, in positing a distinctly affective kind of violence, Whitney is not minimizing other more (overtly) physical forms of violence, but is calling attention to that violence “that functions by quarantining affective force within the racialized body until it becomes toxic” (Whitney 2018, 497). This *violent* toxicity refers to the negative impacts of the kind of mis-metabolized affects mentioned above, and can be quite literally physical, as evidenced by the following passage Whitney draws from Fanon’s *The Wretched of the Earth:* “In the colonial world, the colonized’s affectivity is kept on edge like a running sore flinching from a caustic agent (1968, 19).” The colonized, or the person who experiences affective violence, is harmed affectively, psychosocially, and even biologically, by the affective wronging they face.

Gallegos and Others

As mentioned above, the literature on affective injustice is growing. A more recent paper of Whitney’s, as well as one by Myisha Cherry that considers the dangers of affective stereotyping, are mentioned above. Perhaps the most substantial contribution to the literature that has come after Srinivasan and Whitney’s papers is that of Francisco Gallegos. Gallegos’ writing on affective injustice begins from a lack, namely of a systematic — and, it seems by extension, a singular — account of affective injustice that is less broad than simply an injustice that targets someone’s affective capacities.

Gallegos argues that affective *justice*, whether on Srinivasan’s or Whitney’s account (or Archer & Matheson’s, which I will describe shortly) of the requisite injustice, ought to be defined “as a state in which each person has the *affective goods* they are owed” (2022, 5). Accordingly, then, affective *injustice* refers to “the morally objectionable deprivation of such goods”; Gallegos’

[25] Emotional labor is explained in more depth in the next chapter. Hochschild defines it as that labor which “requires one to induce or suppress feeling in order to sustain the outward countenance that produces the proper state of mind in others” (1983/2012, 7).

footing in the distributive justice paradigm is thus clear (ibid). The affective goods that, in Gallegos' view, ought to be distributed if we are to reach affective justice include *subjective well-being* (vis-a-vis the "pursuit of happiness" [ibid, 12]), and *emotional aptness* (understood as "the fit or harmonious correspondence between evaluative properties in the world and one's emotional response to those properties" [ibid, 8]).

Gallegos acknowledges the limitations of a distributive approach to affective — and, it seems, any form of — injustice, writing that the fundamental affective goods he lists "are not the kinds of goods that can be demanded or directly bestowed or guaranteed" (ibid, 12). However, he maintains that these fundamental goods are bolstered by a series of "subsidiary goods" that actually enable people to pursue those fundamental goods of subjective well-being and emotional aptness (ibid). These subsidiary goods with respect to the *fundamental good* of subjective well-being include:

- "*Affective freedoms*, such as freedom from interference in the pursuit of subjective well-being, including freedom from circumstances that give rise to emotional distress and negative or unpleasant emotions and moods.
- *Affective resources and opportunities*, such as materials, activities, and circumstances that contributive positively to one's subjective well-being, including nurturing interpersonal and social relationships; sleep, therapy, and other means of providing self-care; and 'affective scaffolds' (Maiese 2016)[26] in the build environment that facilitate positive mood and self-evaluation.
- *Affective recognition*, such as respectful consideration of, and responsiveness to, one's particular needs with regard to subjective well-being" (Gallegos 2022, 7).

Even though Gallegos is clear that subjective well-being may not be distributable *in itself*, he suggests that these subsidiary goods can be and are often distributed. In the *disabled rage* case, he would likely say that the injustice lies in someone's being denied each of these goods. For instance, that ableism can produce rage represents a loss with respect to 'affective opportunities,' in that it means the *opportunity* not to be angry has been curtailed.

Among early responses to the original affective injustice papers were Archer and Mills (2019)

[26] In this paper, Maiese argues that research on the relationship between environment and cognition can be extrapolated and used to support similar conclusions about the relationship between environment and *affect*.

and Archer and Matheson (2020). Archer and Mills describe affective injustice as, at its core, a "demand that [victims of affective injustice] regulate away their anger," which entails a claim that their anger is unwarranted to begin with (2019, 79). The pair focus on Srinivasan's original account of affective injustice, arguing that in trying to sort out the normative conflict that they're experiencing, victims are taxed with *regulating* their anger (rather than solely repressing it) (ibid). Archer and Matheson draw from both Srinivasan and Whitney's accounts of affective injustice, focusing on what they call "emotional imperialism": "a powerful group['s] imposing aspects of its culture's emotional norms and standards on another less powerful group whilst at the same time marking out the other culture's emotional norms and standards as deviant and inferior" (2020, 11). The authors' paradigm case of emotional imperialism is what they call "poppy enforcement," or backlash against British persons (namely athletes) who choose not to participate in the tradition of wearing a red poppy to honor British Armed Forces members who died in World War I (ibid, 3).

Other responses to the affective injustice paradigm include 2023 papers by Macalester Bell, Mary Carman, and Lisa Tessman. Bell argues that, insofar as our emotions express our values, utter[27] affective dismissal is akin to "dismissing someone as a valuer" (2023, 241). She argues, further, that affective dismissal constitutes an affective *injustice* — rather than a one-off affective wrong or harm — though she (as I will in the subsequent section and chapter) argues that current models of affective injustice are insufficient. Ultimately, guided by a Rawlsian view of injustice, Bell argues that the injustice (and not merely the wrong) of affective dismissal lies in "the Injustice of Emotional Dismissiveness which arises when a person's or group's emotions are utterly dismissed out of prejudice…[which] hurts society as a whole by depriving all of us of the valuations of some persons or groups" (ibid, 263). Framing some varieties of affective dismissal as unjust, rather than simply wrong, creates a stronger moral prohibition against them.

[27] Bell clarifies that *utter* affective dismissal, as opposed to affective dismissal simpliciter, is when a person is entirely dismissed as a valuer. This echoes Bell's claim that affective dismissal falls along a spectrum, rather than being an all-or-nothing happening (2023, 245).

Carman synthesizes Thaddeus Metz's writings on harmony — that is, promoting the conditions that respect our capacity for harmonious relationships — as an utmost moral obligation with the affective injustice literature. She, like Srinivasan and Whitney, takes "disruptive emotions" as her starting point, arguing that even emotions which appear on their surface to promote *disharmony* might deserve moral weight in a society that values harmony because they promote "other things that we intuitively deem of moral significance, such as human rights" (Carman 2023, 3; 6). That is, Carman cautions us against an overly simplistic view of emotions as pro- or anti-harmony, because even some immediate disharmony can be justified with respect to broader moral norms, including a larger-scale harmony.

Similarly, Tessman argues that some relationships are such that they ask us to step outside of objective apt-or-not-apt judgments. Instead, she writes, "justice requires that we take an intersubjective perspective" and recognize that some value judgments (insofar as emotions reflect evaluations) are not supposed to be shared or universal (Tessman 2023, 206). She cites unconditional love as a paradigmatic case where searching for reasons of fit — of aptness — might lead us to "miss the opportunity to simply witness — and lovingly accept — what is revealed about [the loved] by their full range of attitudes or emotions" (ibid, 211). Tessman's argument, then, amounts to a call to set fittingness conditions aside in some cases because, to put it simply, aptness is not always of utmost importance.

Shortcomings of Affective Injustice

The extant affective injustice literature has undoubtedly been helpful to those of us whose work focuses on the affective dimensions of injustice, and those of us who appreciate yet are unsatisfied by the notion of epistemic injustice. Yet even those theories that are relatively complete or comprehensive possess a few critical shortcomings that prevent them from truly being systematic theories. I focus here on three overarching critiques of extant paradigms: (1) They rely on a sometimes misleading hierarchy of injustices; (2) They leave ambiguous precisely when an infraction — when an affective *wrong* — rises to the level of an injustice, thereby suggesting that all affective harms are unjust; and (3) Most critically, in my view, they are unable to account for injustices that do not begin with *uptake*, and so cannot reckon with the fact that we can act wrongly — we can breach our affective ethical duties — even without anyone knowing. The third critique becomes especially salient, I show, when we consider wrongful emotions like those described in the previous chapter given that such emotions might never be made explicit.

Readers will notice that the critiques leveraged below (taken as a whole) operate on two levels. First, they suggest that the affective injustice literature falls short of its own 'internal' goals. That is, confusion remains about the paradigm's applicability to the full range of phenomena that straddle the emotions and injustice. Second, the below critiques call into question whether affective (in)justice is too narrow a concept, one unable to account for the full range of ways that we can wrong others through or with respect to emotions. This is why I find myself drawn toward speaking about affective ethics, rather than affective injustice. An ideal theory of affective injustice should be internally consistent and should capture a wide range of emotion-facing wrongs grounded in social injustice. However, such a theory is merely a node of a broader affective ethos, one which can guide our dealings with others' — and our own — emotions without requiring too firm a grounding in social injustice.

Ordering Injustice

The extant literature on affective injustices relies on a particular ordering, a stratification, of injustices. This is most obvious in Srinivasan's work; recall from the subsection on her theory that she describes it as a "*second-order injustice* that is parasitic on *first-order injustice*" (2018, 135; emphasis added). The first-order injustices that she is concerned with are those most often identified as pervasive prejudices that can rise to the level of structural injustices: racism, sexism, ableism, and the like. To rightly claim that an affective injustice is taking or has taken place, there must be some identifiable primary injustice constituting its basis. In her central example of Black radicals being dissuaded from anger because of its (alleged) counterproductivity, the first-order injustice is anti-Black racism. The affective injustice can only take place because this racism is present, because the first-order injustice forms the basis of certain groups' expressions (e.g., of anger) being stigmatized more than others', and those groups' thus being subject to stronger admonitions about counterproductivity (Srinivasan 2018, 138).

The focus on prejudice — on first-order injustices forming the base of affective ones — features in Whitney's work too. This was first raised by Macalester Bell, who suggests that the crux of affective injustice is "a person's or group's emotions [being] utterly dismissed out of *prejudice*," resulting in a "deep disrespect for persons as valuers" insofar as emotions are evaluations (Bell 2023, 263; 242). The recourse to prejudice here is a clear reference to Fricker's notion of epistemic injustice as relying on identity-prejudice. Whitney cosigns this in her follow-up to her 2018 paper, writing not of the uptake economy simpliciter but of the "prejudicial uptake economy" (2024, 55). Only when this *prejudicial* uptake economy is reproduced, Whitney writes, is a behavior an example of affective injustice (ibid, 49). Thus, her version of affective injustice too renders it second-order with respect to first-order injustices or prejudices.

I have argued elsewhere (Bigelow 2024) against relegating affective injustice to second-order status, as I disagree with the idea that it can only happen once some other injustice has taken place (or is present). I suggest that affective injustice might actually sometimes be primary, falling ontologically prior to or on par with other injustices (those classified as first-order above). My chief case study is the rhetorical transformation of women's anger into Madness or insanity, a phenomenon discussed at more length in the next chapter. I believe that affective injustice often forms the very *core* of prejudices (or 'first-order injustices') like sexism rather than merely being accidental or subsequent. This is evidenced by sexism's being reinforced (rather than simply emphasized) by the phenomena targeted by affective injustice.

More than this, though, I want to highlight that there are cases where injustice is neither primary nor secondary; cases where it does not figure at all. This is part of the motivation for raising DR2; I am not convinced that the conflict wrought by competing access needs is a matter of justice or injustice, and it seems that neither Jay nor Kay has ‘ableist power’ over the other. I recognize, though, that it can be difficult to conceptualize conflicts about disability as devoid of (or at least as mostly not reliant upon) ableism. The following subsection expands upon this bit of my dissatisfaction with the affective injustice literature and concludes with an example from Iris Murdoch’s work — an example that figures heavily in this dissertation’s fifth chapter — that removes identity from the equation altogether.

Harm without Injustice

Related to the above critique of the affective injustice literature, I find that extant accounts make it difficult to parse exactly which affective harms or wrongs rise to the level of injustices, a considerably higher normative bar to clear. Theorists might reply that they rise to the level of injustices when social injustice is involved. My reading of DR2 shows this to be an unsatisfactory answer. Even more-so, though, I find the affective injustice literature to be too narrow in scope to serve as the basis for an affective ethics, because even equally situated persons can mistreat one another. Even these more mundane interactions, I contend, are worth accounting for. This section, then, is more a meta-critique showing the limitations of affective injustice as a concept.

It may help to ground this in specific examples from the literature. Recall Gallegos’ picture of affective injustice as consisting of the deprivation of affective goods. While it is true that, by allowing for a wide variety of affective goods, his model can account for a wide variety of affective injustices, he may paint in too-broad strokes for the affective injustice literature. For example, his account of subjective well-being entails an entitlement to “freedom from interference in the pursuit of subjective well-being, including freedom from circumstances that give rise to emotional distress and negative or unpleasant emotions and moods” (2022, 7). My disagreement here arises because I do not think that circumstances leading to negative moods — or even those moods themselves —are necessarily unjust, because they need not be grounded in social injustice or prejudice. Emotions like contempt, typically read as negative, can be apt responses to the injustice of the world (much like Srinivasan's apt anger) but, furthermore, they can be guiding lights in shaping a *better* world (see Bell 2013). This presupposes that in an ideal world these emotions would not come about in the first place,[28] but the world we live in is notably non-ideal, and thus contempt (and anger, and other typically negative emotions) need not be emotions that we should (want to) avoid. Thus, I am wary of Gallegos' implication that freedom from negative emotions (or from the experiences that produce them) is necessarily tied to justice. To be sure, some — many, even — circumstances that call for negative emotions are unjust, but a careful account of affective injustice should delineate precisely when experiencing negative emotions is wrongful, when it is harmful, and when it is actually unjust. The extant accounts fail to do this, and so they fail to delineate the ways that affective injustice is but a component of affective ethics.

Similarly, Srinivasan's account focuses on the normative conflict that victims of affective injustice are forced into, but she fails to set forth what it is that makes a *particular* normative conflict unjust (rather than, say, inconvenient or uncomfortable) because her claim is simply that needing to choose between one's emotions and bettering one's standing, when placed against a backdrop of first-order injustice, is sufficient to mark an affective injustice. Bell makes a similar critique, asking why "we should regard this conflict as a new form of injustice," as doing so seems to entail a claim that we ought (deserve, have a right) "to be conflict-free in our affective lives" (2023, 260). This parallels Gallegos' claim mentioned above as well. We are frequently in situations that ask us to choose between expressing our emotions and bettering our situation, but I would hesitate to call the full range of these situations *unjust*. Some are, to be sure, but there are cases where having to choose between displaying a feeling — even an apt one — and bettering your situation might not be unjust. In this way, my central critique of Srinivasan's view of affective injustice might be summed up as follows: she presumes the

[28] To be sure, I am not even convinced by this, by the idea that negative emotions make the world a worse place.

legitimacy, or the rightfulness, of anger, even as she acknowledges that anger can be of greater or lesser "moral standing" (2018, 131).

To illustrate, consider the premise of a childhood favorite book series of mine: Lemony Snicket's *A Series of Unfortunate Events,* which documents the (fictional) trials faced by a group of young siblings who stand to inherit an enormous fortune upon turning eighteen after losing their parents in a mysterious fire. Throughout the thirteen-book series, the children — the Baudelaires — must outwit Count Olaf, the series' villain, in order to protect their fortune from him. If Count Olaf were to openly express his glee at the prospect of scheming three children out of their inheritance in front of, say, the family's estate manager, the estate manager might raise eyebrows. It makes sense, then, to say that Olaf is forced to (however poorly) conceal his happiness at the prospect of stealing the children's wealth. However, I would *not* say that the villain's concealing his glee in order not to tip off the estate manager is an injustice, even though he is in effect forced to choose between expressing his happiness at the prospect of stealing the Baudelaires' fortune and maximizing his actual likelihood of stealing it (since, it seems, if he were to express such happiness right off the bat in the first book the children would not be placed in his custody). We might still draw some attenuated relationship to injustice here — saying, for instance, that Olaf's emotions rest on his being an adult, and drawing upon the emerging literature on philosophy of childhood and children as an oppressed group — so some might argue that this case still meets the narrower criteria of affective injustice.

Another example might be more fitting, then. In "The Idea of Perfection," Iris Murdoch asks readers to consider a woman (M) who "feels hostility to" her adult daughter-in-law (D). The contours of this example as presented by Murdoch will be fleshed out more thoroughly in Chapter 5, but here it might be apt to consider modifications of it. Let us imagine that M and D are similar in all the features typically highlighted in conversations about injustice: race, gender, sexuality, class, ability, and so on. M's maintaining negative emotions toward D, then, might have no grounding in or bearing on injustice. Similarly, if M were to treat D poorly (which, to be sure, she does not in Murdoch's telling), dismissing her emotions (about any and everything) outright and making the younger woman feel that she must walk on eggshells around the older, it is not clear to me that this is an affective injustice, rather than simply a violation of an affective ethics. M might be mistreating D, but this does not mean she is treating her unjustly.

Consequently, I do not find that we need prejudice to play quite so central a role in our account of affective injustice. Making affective injustice reliant on first-order injustices means that members of (relatively) dominant groups cannot experience affective injustice. First, this misses the complexity of prejudice, and how it fails to conform to identity binaries where one 'half' is oppressed and the other is not. Second, relying on prejudice or first-order injustice masks the ways that affective wrongs can exist irrespective of prejudice. Explicitly linking prejudice and affective injustice (qua structural phenomenon) might too cursorily overlook interpersonal (rather than structural and thus broader reaching) injustices by seeking explanation for interpersonal affective injustice sin terms of prejudicial social structures. I do not want to deny even that the presence of prejudice might make an affective injustice *worse*, but I similarly want to maintain that affective injustices can occur even in circumstances *completely* (and not just nominally) unguided by prejudice.

Affective Injustice and Uptake

My most broad-reaching concern about the existing affective injustice literature bears on its ability to account for injustices like those discussed in the previous chapter; unjust *feelings* — like those feelings that maintain the ableist imaginary — themselves. My final concern with the affective injustice literature is that it is unable to account for injustices that are not premised on affective uptake. In Chapter 5, I remark on how championing uptake complicates our ability to see moral development in Murdoch's M. Below, then, I focus more on the ways that the affective injustice literature displays its commitment to uptake.

As mentioned above, both of the dominant theories of affective injustice are based on Miranda Fricker's concept of epistemic injustice. Both testimonial and hermeneutical (epistemic) injustices — the two species of epistemic injustice — have their basis in *uptake*. Testimonial injustice is about deprivations of uptake simpliciter, where an individual is denied uptake because of their being judged uncredible. Hermeneutical injustice, on the other hand, has to do with individuals accessing the kinds of interpretive (hermeneutical) resources that would allow their testimony to receive uptake. Perhaps it is through being modeled after Fricker's epistemic injustice that affective injustice gets its being uptake centered. This is most straightforwardly apparent in Whitney's picture of affective injustice as "the weight afforded [an individual's] feelings" being damaged by these feelings being stripped of their "affective force" (2018, 495). Whether through affective marginalization, exploitation, or violence, affective injustice under this understanding pertains to individuals' feelings not being *taken up* well. The force that an emotion ought to have is either denied altogether or is displaced onto the wrong targets; the latter is the case in Audre Lorde's writings about Black women being treated as "affective waste disposal site[s]" (Whitney 2018, 499).

Srinivasan's picture of affective injustice is uptake-based in a slightly less straightforward way. Her account overlooks cases where there may be no such normative conflict at all because the relevant emotion remains unspoken. For example, a person who bears the ableist emotions discussed in the previous chapter might not actually make these emotions known to any disabled person(s), and they might never intend to do so. I want to maintain that their emotions are unjust despite their never being made explicit, and despite the people bearing these emotions' never having to choose between expressing them and bettering their circumstances. This speaks to a critique, explored more in the subsequent chapter, about accounts of affective injustice relying too heavily on uptake and (because uptake requires expression) on an injustice's being made explicit. This also speaks to a concern about the direction of fit of affective injustice too, because the person who would hypothetically be expressing an emotion is here not the target of the injustice but is instead its perpetrator.

The very nature of the normative conflict Srinivasan characterizes as central to affective injustice is such that it depends on either actual or hypothetical poor uptake. That is, victims of oppression would not find themselves bound by this normative conflict were they not armed with expectations that members of dominant social groups would take their emotions up in ways that harm them. Myisha Cherry explores the 'Angry Black Woman' stereotype in the context of Srinivasan's view of affective injustice. Did Black women not have a well-founded expectation that their emotions would not be taken up justly, they would not experience this conflict (and, by extension, affective injustice).

Whitney's follow-up to her 2018 paper also acknowledges a flaw in her original picture of affective injustice: she fails to explain what makes a particular failure of affective uptake unjust. Bell explains this in terms of Whitney's lacking a positive account; she fails to provide an account of affective *justice* by leaving ambiguous whether we have an all-things-considered right to affective uptake. Bell argues, then, that "it is unclear how we could distinguish being properly moved by another person's emotions and being manipulated or coerced" by them according to Whitney's 2018 theory (Bell 2023, 259). Whitney acknowledges this shortcoming, noting in her 2024 paper that "injury is not a sufficient condition for injustice" (53). In other words, being denied uptake, and even being harmed *by* being denied affective uptake, cannot be sufficient for affective injustice. There needs to be some further criteria to determine when these injuries rise to the level of injustices.

This is a practical shortcoming rather than merely an oversight for the reason that Bell identifies: it might inadvertently license manipulation or coercion. Whitney's account might, then, allow for malicious claims that a person is owed more uptake than they are receiving. In the cases described in the previous chapter, what is to stop the ableist person — the one who adopts the attitudes of the ableist imaginary — from claiming that their fear, pity, and awe are not receiving their deserved uptake? Without a clear delineation of when (and which) affect(s) deserve uptakes, others too become vulnerable to being compelled to engage with unjust affects like those in the previous chapter.

An account that relies too heavily on uptake cannot contend with the injustice of the attitudes described in the previous chapter. These attitudes do not derive their unjust-ness from their uptake; they are unjust, I propose, even if they remain unarticulated. If they remain unarticulated or unexpressed, they are decidedly not even *offered* uptake. Thus, if we remain committed to an uptake-based account of affective injustice, the paradigm seems unable to reckon with the injustices introduced in Chapter 1.

Conclusion

In what precedes, I provided an overview of the affective injustice paradigm as one of the foremost attempts to systematize an account of what we owe to others' emotions. The most fully fleshed-out accounts of affective injustice are those of Amia Srinivasan, Shiloh Whitney, and Francisco Gallegos, though the literature's rapid growth in recent years has led to emerging accounts and descriptions of cases that should fall under the affective injustice umbrella. The affective injustice literature may have the power to alter the landscape of feminist philosophy in similar ways to Miranda Fricker's work on epistemic injustice and the literature that grew out of it, given that it seems to track a phenomenon that is nearly ubiquitous. Shedding light on the different ways that affective injustice materializes can likewise draw attention to the harms that it causes, thus creating stronger momentum for rectifying such injustices. This work is particularly important given the banal view of philosophers as champions of reason rather than the emotions. Taking emotions seriously as matters of justice is a step, then, toward recognizing that we can treat others better or worse with respect to their emotions.

Affective injustice is a useful category but does not provide a comprehensive affective ethics. Even if these accounts make it clear that the targeted phenomena are injustices, they do not explain why they are best described as injustices are not, for instance, wrongs or harms. This makes it difficult to derive a positive notion of affective justice. If it is not abundantly clear what makes an injustice an injustice, how are we to avoid perpetrating injustice? Moreover, if injustices are never brought to light — as in cases where the person bearing unjust attitudes never expresses them — how are we to work toward repairing them? The other major weakness of the affective injustice literature is — perhaps paradoxically given the last few sentences — that it might rely too heavily on injustice as a category, thereby suggesting that the only affective *wrongs* are affective injustices. An ethics that focuses primarily on injustices captures something important, to be sure, but it remains incomplete because the ways we wrong one another often do not rise to the level of injustice.

Chapter 3 - Marginalization, Psychiatrization, and Emotion Norms

Introduction

The two previous chapters have showcased some of the conditions making an affective ethics necessary, along with philosophers' attempts at systematically crafting such an ethics through the affective injustice paradigm. In this chapter, I explore another type of harm that comes about with respect to emotions, but that I believe is left out of current conversations about affective injustice because of their focus on uptake as mentioned in the previous chapter, and because of their focus on emotions straightforwardly (or simpliciter) as discussed below. Rather than looking at the mistreatment of individuals' discrete emotions, or emotions that are fundamentally rooted in broader social injustice, I study here the misapplication of expectations *relating to* emotions, particularly social norms that prescribe certain behaviors in response to others' emotions. Importantly, insofar as they are taken up as general social norms, these norms are often enacted not in cases of social injustice, but in more mundane dealings with others qua affective beings.

This chapter proceeds as follows: I begin by identifying two distinct registers of emotion norms, emotion norms *simpliciter* and not-strictly-emotional social norms that arise adjacent to emotions, referred to as *emotion-responsive norms*, or simply *responsive norms*, for the sake of brevity. Emotion norms simpliciter dictate which emotions are appropriate when, and for or expressed by whom. Their targets are thus (individuals') emotions themselves. Responsive norms, on the other hand, do not dictate the aptness or normativity of particular emotions, but instead determine the range of appropriate responses to emotions (typically, emotions expressed by others). Such responses are linguistic and behavioral, and they more closely accord with ‘social norms’ as commonly understood. After setting forth this distinction in norm types, I begin exploring the ways that emotions come to be pathologized and, more specifically, psychiatrized via the application of norms pertaining to when and in what magnitude a particular emotion is appropriate. An example I consider here is the transformation, rhetorical and otherwise, of oppressed groups’ righteous anger (especially anger that is about oppression) into Madness or insanity. This is a similar injustice to the affective injustices discussed in the previous chapter because it pertains to emotions themselves being treated poorly. Both kinds of emotion-norms — those that directly target emotions and responsive ones — are complicit in the creation of (a population that Sara Ahmed terms) *affect aliens*.

I then suggest that the uptake-based nature of affective injustice as described in the “Shortcomings” section of the previous chapter renders the paradigm unable to account for misapplied emotion-responsive norms, that is, harms that stem not from emotions themselves but from emotion-adjacent behaviors. This is where I introduce ‘therapy-speak’ as a specific kind of psychiatrization, but one that merits separate consideration from the discursive transformation mentioned above. I maintain that the affective injustice literature might be unable to ‘handle’ or account for responsive norms (including and especially the species of psychiatrization operative in ‘therapy-speak’) when they occur not between oppressed and oppressive parties, but among peers; recall the previous chapter’s discussion of the ordering of injustices (e.g., affective injustice as secondary to [primary] social injustices) and affective injustice theorists’ unwillingness to establish a hard line between harmful behaviors and unjust ones.

Two Registers of Emotion Norms

Emotion norms can be thought of on two levels, differentiated by their proximity to emotions themselves. The first register of emotion norms consists of emotion norms *simpliciter*. These are the kinds of norms that Vishkin and Tamir identify as expressing “which emotions are or are not experienced or expressed, and which emotions should or should not be experienced or expressed" (2023, 453). This definition is reinforced especially by Kurth's understanding of emotion norms as "concerning (e.g.) when an emotion is appropriate and how to respond to inappropriate[29] emotions" (2022, 3). Emotion norms tell us when to feel what and, adding the second element of Kurth's definition, they set up the punitive scripts for those who fail to feel correctly by dictating norms of (for instance) disgust or anger at those whose emotions transgress norms too boldly, those who fail to feel correctly. Sociologist Arlie Hochschild also identifies emotion norms as what she terms *feeling rules*. Feeling rules are essentially emotional scripts or conventions, identified "by inspecting how we assess our feelings, how other people assess our emotional display, and by sanctions issuing from ourselves and from them," according with Kurth's definition (Hochschild 1983, 57). Feeling rules or (first-order) emotion norms simpliciter are not absolute or universal; different social groups might and often do have different allowances for particular emotions, but what remains is that "there are rules or norms according to which feelings may be judged appropriate [with respect to] to accompanying events" regardless of what the actual content of these norms dictates with respect to a particular event (ibid, 59).

Some of these emotion norms are fairly straightforward, and making sense of them does not require recourse to different axes of marginalization and oppression. These are the more straightforward norms about how to feel, or at least what feelings to express, in a certain situation. Much of this is conveyed without ever needing explicit teaching. Joy and giddiness might be appropriate at an athletic event or an awards ceremony but are markedly inappropriate at most funerals and memorial services. Sobs of despair, likewise, are thought unfitting in response to a friend who has shared news of their engagement, or to a grocery store running out of your preferred brand of potato chips. To use one of Hochschild's examples, a flight attendant rolling their eyes or erupting into a fit of rage every time that a passenger made a simple request

[29] Kurth's focus on inappropriate emotions here is important for distinguishing his focus from the responsive norms that I am concerned with later in this section.

would likewise be thought inappropriate (see Hochschild 1983, 93). Just as an emotion can be judged as wrong irrespective of its relationship to social injustice, it is worth noting that we can *wrong others* by holding them to particular emotional expectations irrespective of injustice. In other words, we can treat others' emotions poorly even absent structural injustice because, for instance, we can dismiss them out-of-hand. Phenomena like gaslighting, which undermine individuals' emotional capacities, often occur between relative equals in the kinds of relationships we think of (or at least hope to think of) as egalitarian (see Abramson 2024).

Some emotion norms *are* closely related to oppression and thus injustice, though. The two examples that I consider here pertain to gendered and racialized oppression. Alison Jaggar points to an example in her writing about the supposed conflict between reason and emotion (and feminist emotions as a corrective to this conflict) in contemporary philosophers' work. Feminist emotions, writes Jaggar, are those that "incorporate feminist perceptions and values," as anger becomes feminist "when it involves the perception that the persistent importuning endured by one woman is a single instance of a wide-spread pattern of sexual harassment" (1989, 167). In other words, anger at a particular act of gender-based violence may well be *feminist* anger because it reflects the broader ethos of the social movement. She identifies feminist emotions like this anger as "outlaw emotions" that are "[incompatible] with the dominant perceptions and values," thus rendering them "conventionally unacceptable" (ibid). As explained in a subsequent section, emotions being deemed unacceptable creates *affect aliens*. This can easily be recapitulated in terms of emotion-norms; there is a norm against feminist anger because, to speak in broad terms, this anger does not "serve the interests of [the] rich white men" who benefit from the patriarchal forms that feminists push back against (ibid, 165). Emotion-norms do not just dictate when and for whom an emotion is improper, though; they can lead to negative outcomes for those who express such feelings. This can look like the negative consequences (alienating would-be allies, losing social position, inciting further oppression, and so forth) that Srinivasan touches upon in discussing the counterproductivity critique (of oppressed groups' righteous anger). A common exemplar here is women in politics often being warned against seeming too angry (and women's anger being read as the gendered *shrillness*) when the same anger is thought to be becoming of male politicians, marking them as strong leaders. The consequences might also consist of (further) social and political marginalization, making one's political goals less likely to come to fruition. Thus, feeling incorrectly can meet the charge of counterproductivity.

Some of the emotion norms that weigh especially heavily on members of marginalized groups operate on the basis of race rather than gender. As described by both Srinivasan's and Whitney's models of affective injustice in the previous chapter, Black people's anger is often policed in ways that white people's anger is not. From this we can infer a norm, and quite a powerful one, depicting Black people's anger as less apt — as less socially acceptable an emotion — than white people's. This norm is enforced, for instance, through responses like guilt; consider Audre Lorde's writings on Black feminists' need to contend with white feminists' guilt upon being presented with evidence of anti-Blackness in the feminist movement (1984). This guilt tells Black women that their anger is inapt and uncalled for, even when it is perfectly justified. It is *incompatible* (to use Jaggar's term) with the way that Lorde's (white) interlocutors see themselves and is accordingly met with resistance.

As mentioned, emotion norms simpliciter are not the only kinds of norms at play in this chapter. I remain chiefly concerned with a second set of norms that not only dictate how we (and others) ought to feel, but *how we ought to respond to others' feelings*. These are different from the norms described in the previous paragraph, which take the simple emotion norms as their object of inquiry; simple emotion norms allow us to judge others' emotions for their aptness, but still do not say much about how we are to construct the kinds of schemata that allow us to decipher when an emotion is (in)apt. The distinction between registers of norms boils down to this: emotion norms simpliciter are about how we feel, whereas the second register of norms are not about how we feel but what we say or do in response to others' feelings.

Because they do not have to do with whether a particular emotion is appropriate, they are not, properly speaking, emotion norms. They are instead more akin to social norms. I refer to them as emotion-responsive norms, or responsive norms as a shorthand, to distinguish them from emotion norms simpliciter. These norms are the target of much of this chapter's remaining analysis, so when I speak simply about *norms* — without specifying emotion norms *simpliciter* — these responsive norms are what I am talking about.

The responsive norms I am concerned with are not quite the same as what Hochschild terms *emotional labor*, which she defines as "the management of feeling to create a publicly observable facial and bodily display" (1983, 7). Her paradigm case of emotional labor is that performed by flight attendants, whose "*emotional style of offering service is part of the service itself*"*;* in other words, a flight attendant whose personality seemed cold and detached to their passengers would seem to have failed *at the task of being a flight attendant*, and not simply at some arbitrarily set company policies about passenger treatment (ibid, 5; emphasis in original). Much of emotional labor consists of emotion management or regulation. Importantly, though, Hochschild seems concerned only with a person's management of their *own* emotions, albeit for the sake of regulating others' emotions. For example, a flight attendant in an emergency might be tasked with managing passengers' fears, but the emotional labor paradigm points to the ways that they do this through regulating their own emotions and the expression(s) thereof. The norms I am concerned with might be emotion-management strategies, to be sure, but they do not actually target feelings. They target, instead, actions including verbal ones, particularly those actions that arise adjacent to others' emotional expressions. Though those who hold jobs requiring a lot of emotional labor are, indeed, often charged with managing others' emotions (consider flight attendants' obligation to keep passengers calm during an emergency landing), this does not seem to be what Hochschild was targeting with her discussion of emotional labor. Emotional labor pertains more to an individual's regulating their *own* emotions, rather than regulating or managing their responses to others' emotions.

Marginalization through Psychiatrization

In this section, I discuss the psychiatrization of (already, otherwise) marginalized persons' emotions as a distinct form of emotional marginalization. Psychiatrization often occurs in response to (those who transgress) emotion norms. My paradigm cases here involve righteous anger borne by women and racial minorities (particularly Black people) — the two groups whose outlaw emotions are discussed above — being (mis)read as Madness[30] or insanity. This, I contend, is an instance of first-order emotion norms at work, and I show below that the extant affective injustice literature can account for it quite well.

Recall from the previous chapter Whitney's usage of *marginalization* — one of Young's five faces of oppression — to explain one layer of affective injustice. Young understands *marginals* (the marginalized) as "people the system of labor cannot or will not use," and who are thus "expelled from useful participated in social life and thus potentially subjected to severe material deprivation and even extermination" (1990/2011, 53). Because the contexts I am concerned with are not strictly labor-related (or at least are not always or explicitly so), it makes sense to understand marginalization in affective contexts as exclusion or expulsion more generally, particularly from membership in some (often ill-defined) affective community.[31] Whitney defines affective marginalization as "an expulsion from participation in affect circulation that depletes affective agency, influence, or authority" resulting from persons' emotions being denied uptake (ibid, 497). Marginalization can appear differently too, though. It is not always an outright or complete refusal of uptake, or a dismissal of individuals as having affective any capacities (see Whitney 2018, 506; Whitney 2024). Someone can be marginalized by virtue of their emotions being depicted as less-than-important, even if the basis of this is more one-off dislike rather than systemic injustice.

[30] I use 'Madness' here in the same way that those in the Mad Studies tradition do. It points to the phenomena otherwise classified as insanity or mental illness. I capitalize the 'M' in Mad/Madness to distinguish it from uses of madness-as-anger.

[31] I understand "affective communities" as loosely constrained groups beholden to the same affective norms, particularly first-order emotion norms.

I posit that marginalization can also occur when affective energies are warped, or transformed into something they are not, because dominant frameworks for understanding emotions are inadequate. Whereas Whitney focuses on a single emotion's needing to be 'metabolized' by someone other than its original bearer, I am concerned with emotions that become substantially transformed into new affects. As explained more below, the cases I have in mind are ones where a distinct affect — anger, say — is taken up as something else entirely. The cases Whitney is concerned with, on the other hand, have more to do with (e.g.) anger being taken up *as anger* but having its effects (again, *as* the effects of *anger*) felt by a different party than the original angry person. There is no transformation or change here. Both these cases play on first-order emotion norms, but they do so in different ways.

Lavallee and Gagné-Julien make a similar point, arguing that psychiatrization — a type of marginalization to be defined shortly — mirrors Dotson's "contributory injustice" or Pohlhaus' "willful hermeneutical ignorance," additions to Miranda Fricker's original relatively limited idea of hermeneutical injustice (Dotson 2012; Pohlhaus 2012; Lavallee & Gagné-Julien 2024, 17). Interestingly, though Whitney draws most explicitly from the notion of marginalization, the phenomenon she identifies also resonates with Young's idea of *cultural imperialism*, another of oppression's five faces. Young understands cultural imperialism as the experience through which "the dominant meanings of a society render the particular perspective of one's own [marginalized, non-dominant] group invisible at the same time as they stereotype one's group and mark it out as the Other" (1990/2011, 58-9). Non-dominant perspectives are simultaneously invisibilized or erased and rendered hyper-visible. This might be viewed as an extension of or even a precursor to marginalization, because to be pushed to the margins and expelled from social life implies that one's (or one's group's) ways of being are inferior to those of the dominant group. In marginalizing the inferior group, the dominant group others them. Marginalization and cultural imperialism thus complement one another, because marginalization presupposes a sort of stratification of cultures.

One form of marginalization identified by Lavallee and Gagné-Julien is *psychiatrization*. As the word suggests, this is a kind of marginalization that operates not by pushing individuals outside of public life altogether, but by relegating them to a specific domain: that of psychiatry. This can be done literally (e.g., via institutionalization) or discursively. Discursive psychiatrization can be most simply understood as the description of phenomena (behaviors, discourses, etc.) that are not properly in the domain of psychiatry in psychiatric terms. Beeker et al. attribute the term "psychiatrization," which they characterize as "notoriously hard to define," to Dušan Kecmanovic (1983), who uses it to refer to "psychiatric labeling of social phenomena or of deviance from existing norms" (2021, 3). They define psychiatrization as "a *complex*[32] *process of interaction between individuals, society, and psychiatry* through which psychiatric institutions, knowledge, and practices affect an increasing number of people, shape more and more areas of life, and further psychiatry's importance in society as a whole" (ibid, emphasis in original). Lavallee and Gagné-Julien classify psychiatrization — itself a branch of a broader medicalization — as "neither inherently good nor bad," recognizing the legitimizing role that certain instances of psychiatrization (e.g., the identification of post-partum depression) play (2024, 2; Fricker 2007). What the authors are concerned with, though, is not psychiatrization in general, but the psychiatrization of people's emotional lives. What each of these definitions of psychiatrization has in common is their focus on phenomena being classified as psychiatric that are not, typically speaking, within the domain of psychiatry.

The psychiatrization of our affective lives is evident in "the increasing influence of psychiatric conceptualizations on the ways we interpret our own and others' *emotions*," particularly when those emotions are viewed as abnormal (Lavallee & Gagné-Julien 2024, 3). An apt historical example lies in the uptake reception of women's anger as hysteria, itself once classified in the *DSM-II* as a mental disorder (see Chemaly 2018). In 1989, Elizabeth Spelman

[32] Psychiatrization is complex in part because it can emanate from the top-level (e.g., from clinicians) or from the bottom-up (e.g., from laypersons whose involvement in behavioral health is limited to the patient role).

writes that “anything resembling anger,” when demonstrated by a member of a subordinate social group, “is likely to be reduced to hysteria or rage” rather than being recognized as (justified, apt) anger. Other affects, too, are historically classified as hysteria; Pismenny et al. (2024), mentioned in the previous chapter for their taxonomy of emotional injustices, use Freud’s patient Dora’s discomfort with her parents’ friend making sexual advances toward her as a young girl as an example. Another fitting example is the emergence in the nineteenth century of “drapetomania” as a term used to (further) pathologize via pseudo-medical discourses enslaved persons who tried to escape captivity (Bynum 2000). Other diagnoses related to Black resistance (to enslavement and, post-Emancipation, to racism more generally) have historically included dementia praecox[33] dysaesthesia aethiopis, and, eventually, schizophrenia (Metzl 2009, 28-30). Metzl contends, importantly, that schizophrenia remains over-diagnosed among Black patients in part “because racial tensions are structured into clinical interactions” irrespective of the specific agents (doctors, patients, etc.) involved (ibid, xi; xviii). Moreover, he writes that the *DSM-II* with its rewriting of the diagnostic criteria[34] for schizophrenia “functioned as an *implicitly* racist text because it…enabled users to knowingly or unknowingly pathologize protest as mental illness” upon its release in the 1960s amid the civil rights era (ibid, 98).

This “protest psychosis” phenomenon identified by Metzl may seem fundamentally different from the hysteria example mentioned above. Whereas hysteria seems to be explicitly about pathologizing (women’s) *emotions*, the phenomena Metzl discusses might be classified as primarily about pathologizing resistance. In other words, one might argue that what’s at issue isn’t Black persons’ feelings, but their behaviors’ being at odds with the social position the dominant group wants to relegate them to, regardless of what they are feeling. Though it is

[33] Metzl writes: “Psychiatric authors [of the 20th century] combined Cartwright’s drapetomania with Kraepelin’s praecox in order to contend that African Americans were psychologically unfit for freedom” (2009, 30).

[34] These new criteria included characterizations of schizophrenic patients as having “*delusions, anger, hostility* and *projection*”; Metzl details the ways that these were projected onto Black patients (2009, 101).

undeniably true that diagnoses like those mentioned above are linked with behavior on the surface-level, they can easily be tied to emotions when we consider that behaviors are often (if not always) motivated by *something*. That something is likely a feeling, even if it is one that a person never makes explicit.

Kurth identifies concerns like these as indicative of problematic emotion norms. Emotion norms are “norms regarding when and how it is appropriate (in some sense of appropriate) to experience a particular emotion” (2022, 3). Such norms are typically not wholesale claims about whether an emotion is good or bad; they are instead about the circumstances under which an emotion is apt or fitting. Importantly, Kurth notes that emotion norms are socially constructed and so end up “reflect[ing] the beliefs and values of the socially and politically dominant group; and for this reason, the beliefs and values of the dominant group will tend to be the ones that shape what individuals feel and how they assess those feelings” (ibid). He identifies three distinct worries that feminist philosophers raise about such norms: (1) they force subjects into hard choices, (2) they target the marginalized, rendering them vulnerable to further criticism, which seems to violate standards of justice, and (3) they discount marginalized persons’ agency by claiming that “they’re not doing what they ought to do as a self-respecting, autonomous individual” (ibid, 4-6).

The affective injustice literature is responsive to these three harms. Srinivasan’s picture of affective injustice explicitly — definitionally, even — speaks to cases where subjects are forced into difficult choices. For her, the paradigm ‘hard choice’ is between bettering one’s situation and vocally identifying injustice where it occurs. Whitney’s account, moreover, makes clear the wrong of targeting the marginalized because of her focus on the affective economy. Just as Kurth describes the second criticism as being about cases where marginalization compounds, she describes cases where affective injustice creates greater affective labor for those who are already/otherwise marginalized; consider, here, her use of Audre Lorde’s writings. Lastly, the extant accounts of affective injustice (as I read them) are concerned with cases where the oppressed are denied proper affective agency. Whether this agency is about the ability to determine one's own actions, or the ability to have one's voice properly taken up, or about receiving one's share of the affective goods that allow emotional agency to flourish, existing accounts of affective injustice are concerned with the creation and maintenance of affective agents. In the following section, I take up the example of therapy-speak as a form of psychiatrization that the current affective injustice literature is unresponsive to owing to the literature's dealing solely with first-order emotion norms. I define therapy-speak and the ways that it appears in popular discourse before postulating that it functions as a responsive norm because it governs not emotions themselves but behaviors tangential to emotions.

Therapy-Speak

The type of psychiatrization mentioned above is easily explainable through the affective injustice paradigm. The previous section takes issue with norms that dictate how (when, etc.) we ought to feel particular emotions. These are first-order emotion norms; if women's anger is coherent only as hysteria, there is a prohibition or norm against women's (feeling) anger. The phenomena I examine in this section should then be classified as responsive norms because they pertain not to the feelings themselves, but to the nexus of behaviors surrounding emotional expressions. In other words, I am concerned not with norms dictating how we ought to feel and when, but norms dictating how we ought to respond (not only affectively but also verbally and behaviorally) to others' emotions. This section begins by providing a working definition of therapy-speak, itself a difficult task because the concept is often referred to without being explicitly defined, a kind of 'you know it when you see it' phenomenon. I then argue that therapy-speak often dictates responsive norms because its uncritical application might leave individuals with the impression that others' emotions are always to be granted deference. Moreover, that therapy-speak is so often used in friendly contexts — that it is used *among friends,* and that we tend not to think of friendships as places where social injustice plays out — tells us that it can circumvent any more stringent requirements of injustice.

Recall from the previous section Lavallee and Gagné-Julien's pointing to psychiatrization as the mis- or over-application of concepts imported from psychological discourse. They write:

> The spread of psychiatric concepts into informal use in everyday life is exemplified for instance when we use psychiatric terminology to refer to experiences of ordinary life, such as when talking about someone[35] very organized as 'OCD' (referring to the diagnosis of 'obsessive-compulsive disorder', see e.g., Spencer & Carel, 2021) or when designating all kinds of bad experiences as 'traumatic' (Beeker et al., 2021) (Lavallee and Gagné-Julien 2024, 2).

This kind of slippage of psychiatric diagnoses and terminology into non-clinical conversational contexts exemplifies what has come to be known as 'therapy-speak.' Therapy-speak is a slippery concept that has arisen in public discourse in recent years. Some use it primarily with scare-quotes, because even in espousing nominally therapeutic values, therapy-speak might remain quite starkly separated from actual (psycho)therapeutic practice; it may, in fact, rely on and amplify insights that most clinicians would balk at. Stern and Brackett point to a phenomenon like the one Lavallee and Gagné-Julien note, describing cases of "'therapy' words" where specific diagnoses, symptoms, and psychologically-relevant phenomena — trauma, gaslighting, narcissism, depressed — are misused in a "fuzzy" way where the lines between their proper uses and some (even more) normatively-loaded usage is blurred (2023). There are less pointed, less clinical kinds of therapy-speak too; Amy Franks and David Caldwell cite as therapy speak terms and phrases like "unpack," "creating space," and "sitting with our discomfort," none of which are tied to a particular psychiatric condition (2023, 294). Martha Alexander adds examples like "doing the work," "having the capacity to invest" in a relationship, and talk of interpersonal boundaries (2023). These less diagnostic uses and cases of therapy-speak are the ones that I find especially pernicious because they are more difficult to track. Explicitly clinical language like diagnoses is easier to identify as instances of psychiatrization; a phrase is either used as a diagnostic, or it is not. On the other hand, if the language used is 'psychiatry-ish,' as I believe

[35] It seems that Lavallee and Gagné-Julien's example rests on someone *who does not actually have obsessive-compulsive disorder* describing themselves or being described as '*so* OCD.'

even less clear-cut therapy-speak is, there can be more plausible deniability that a person is engaging in psychiatrization and that their actions should be subject to critiques like those raised by Lavallee and Gagné-Julien.

Because it has only recently come to be explicitly referred to as such, there is a notable dearth of academic literature on therapy-speak. This makes it difficult to land on any definition of it. In an episode of NPR's *All Things Considered*, Rebecca Fishbein describes therapy-speak as the "generally formal" sort of "prescriptive language describing certain psychological concepts and behaviors" (Limbong, 2023). Moreover, and perhaps more importantly for this project, Fishbein claims that therapy-speak is "scripted in a way that removes culpability," presumably analogously to the way that therapy (i.e., psychotherapy) seeks explanation and sense-making without necessarily laying any moral blame on the individual service user (2023). That is, this kind of therapy-speak is crafted in a way that suggests that emotions are not the kinds of things individuals should be blamed for, because it paints emotions as both out of individuals' control and as evaluatively neutral. This is parallel to how in actual therapeutic encounters clinicians are expected to recognize clients' emotions without laying blame on them for negative feelings. In genuine therapeutic contexts, this is oftentimes beneficial. In landing on my own working definition of therapy-speak, I think that the prescriptivism identified by Fishbein is of utmost importance. Thus, I propose defining therapy-speak as such:

> **Therapy-speak:** The carrying-over of concepts from (or purported to be from) psychology into everyday, non-clinical life. It tends to impart a sort of distance between those who use it and its targets, or those it is used in conversation with. Therapy speak can be either diagnostic (as in the examples mentioned by Lavallee and Gagné-Julien) or restorative (as in the non-diagnostic examples described above).

As mentioned, I am primarily interested in the latter. While we might understand diagnostic therapy-speak as the mere misapplication of psychiatric terms, restorative therapy-speak can be understood as follows:

> **Restorative therapy-speak (RTS):** RTS can be defined as emotion-directed

> platitudes that are designed to discharge emotions of their affective force, in part by portraying them as outside of individuals' control and thus not as phenomena that should be 'fed'. A subset of therapy-speak as defined above, RTS often appears as series of emotional DOs and DON'Ts; DO heal your inner child, DON'T disrespect my boundaries, DO hold space for your emotions, DON'T apologize for they ways you're feeling, DO recognize my and your feelings as 'valid' and practice self-care.

This is not to say that all instances of RTS are inherently wrongful. The warnings and encouragements encompassed by RTS can be empowering under some circumstances. Psychiatrist Jessi Gold notes that therapy-speak, particularly because of its prevalence in informal settings (including, as the next chapter explores, on social media), can be a way of making therapy and its discourses more accessible, a noble cause particularly when actual psychiatric services tend to be concentrated in higher-income countries and areas within countries (Blum 2023). Gold notes, too, that these terms can destigmatize mental illness, perhaps encouraging more people to seek care (see chapter 1 for more on disability stigma). What I am concerned with, though, is the prescriptivism such terms are laden with. They are presented not merely as ways that one person is choosing to respond to their own or others' emotions, but as ways that we all *ought* to respond to emotions; as responsive norms. They represent not merely emotion norms or 'feeling rules,' but norms about which feeling rules we ought to feel beholden to; thus arises the distinction between first-order emotion norms and responsive norms because these sentiments become normalized to such an extent that they become compulsory. This is a distinction that I will now elucidate further.

Therapy-Speak and Emotion-Responsive Norms

What is important to establish now is that therapy speak — more specifically the RTS described above — reinforces responsive norms. That RTS reinforces emotion norms simpliciter is clear: when there is an imperative to (or not to) feel a certain way, there is an emotion norm at play. Consider some of the examples from the definition above. "Heal your inner child," "hold space for your emotions," and "don't apologize for your feelings" all dictate norms about how a person should feel with respect to their (own) emotions; that they should not, for instance, feel guilt for allowing their emotions to guide their behaviors. What is a bit less apparent is that therapy-speak (or, as I will argue, the culture surrounding it) also involves responsive norms. This is what I will now discuss. To establish the scope of my inquiry: I will be considering therapy-speak in a dialogical sense, that is, when the speaker and the recipient of the phrase are two separate individuals, rather than looking at self-directed cases. My primary claim in this subsection is that therapy-speak and the implicit cultural attitudes surrounding its use *set* or *maintain* standards for how we ought to respond to others' emotions.

First, some instances of therapy-speak can easily be transformed into the second person and so straightforwardly become guidance on how to deal with others' emotions. "Heal your inner child" is not just absorbed by the speaker with respect to *their own* inner self; it is explicitly a piece of advice given to others about how to deal with their emotions. This phrase can then be translated into something like "you [as in, not-me, as in, not-the-speaker] should heal *your own* inner child." This expectation (that the phrase is as much about guiding others' behavior as it is about guiding the recipient's own) comes naturally, because therapy-speak and similar turns of phrase are necessarily (per the definition reiterated above) not context-specific. Because they function as 'one size fits all' phrases, they must apply to all persons and not only to those who they are explicitly offered to. These utterances thus reflect norms about the treatment we are obliged to give others' emotions, even if they are often presented as guidance about dealing with our own emotions.

Second, even without the explicit translation into second-person language mentioned above, we can draw the connection between therapy-speak adages and second-person moral guidance by drawing upon the notion of *empathy*. As mentioned above, the poignancy of social norms (particularly when those norms occur in emotional contexts, even if they are not properly *emotion* norms) depends in part on empathy or our ability to see others as experiencing emotions like (or at least as) we do. That is, empathy understood in a broad sense does not require that we feel exactly as others feel, but only that we recognize others *as* feeling (that is, affective) beings, much in the same way that we see ourselves. From here, there is a natural leap to assuming that we owe *something* to others' emotions (provided that we want them to feel some obligations toward ours). Thus, even without responsive norms being spoken explicitly via RTS, we can glean that we might have some obligations toward others' emotions. Even if the content of such obligations remains unspoken, the general cultural acceptance of RTS sentiments means that such ideas seem like natural extensions of our affective obligations.

Third and finally, irrespective of particular utterances, therapy-speak reflects a general attitude (one which is amenable to first-, second- and third-person statements) toward emotions. That such phrases as "your feelings are valid" — not in the imperative form mentioned in the definition of RTS but nonetheless often used as an instance of therapy-speak — can be applied uncritically, that is, without appreciation or understanding of the specific contours of a situation, means that they reflect some broader theory of or about emotions. Together such phrases can be read as reflecting a systematic belief that emotions are out of individuals' control and that, as a result, we ought not be critical of them. I understand this as RTS's suggesting that emotions cannot be normatively evaluated. RTS might recognize that emotions can be *instrumentally normatively evaluated* (for how well they serve a particular purpose) or that they can be evaluated for specific fittingness criteria like aptness (see Srinivasan's and Cherry's work mentioned in the previous chapter) but that they cannot be wholesale better or worse. Though this — the idea that emotions cannot be morally evaluated — is a position I fundamentally disagree with as outlined in the introduction to and first chapter in this dissertation, my purpose here is not to leverage this criticism but instead to show that RTS reflects a general attitude toward the emotions. This general attitude, in turn, dictates what we can rightfully be expected to do for others' emotions. If emotions are not normatively evaluable, then we should treat them as evaluatively equal; we should not treat one emotion (of someone else's) as altogether better or worse than another.

Affect Alien-ation: Emotion Norms and Neurodivergence

Although emotion-responsive norms like those promoted through therapy-speak may strike us as natural because of their ubiquity, they are created and recreated as they are instantiated. In this section, I show that the use of therapy-speak as dictating responsive norms contributes to the complex social processes rendering some people — including, most prominently, some neurodivergent and otherwise 'mentally disabled'[36] (to use Margaret Price's terminology [2011]) people — *affect aliens*. I first explain what affect alien-ation is, drawing from Sara Ahmed's body of work, before explaining how this phenomenon is enacted against neurodivergent people in particular, and how it compounds in the pernicious stereotype of the mentally disabled subject as lacking empathy.

Ahmed identifies alienation with respect to the emotions; when happiness becomes linked (causally, discursively, etc.) to certain objects (e.g. the wedding day as the "happiest day of [one's] life," those who experience some other emotion(s) in response to the same objects are alienated or "out of line with an affective community," namely the community to whom (following this same example) wedding days and happiness are aligned (Ahmed 2010, 29; 41). Another example given by Hochschild is of a funeral; "it is possible for a griever to misgrieve," she writes, when they fail to feel (e.g.) sad at a funeral (1983, 63). There are other norms governing grief too, not just whether a griever should be sad but how sad they should be, for how long, where they should express their grief, and so forth. Whatever is felt, provided that it is not aligned with whatever is *supposed* to be felt, is rendered inappropriate (Ahmed 2010, 41; Hochschild 1983, ch. 4 [cf. 59]). Similarly, Ahmed writes, "we can feel alienated by forms of happiness that we think are inappropriate" (ibid, 42). In short, "affect aliens are those who do

[36] Of mental disability, Price writes that: "this term can include not only madness, but also cognitive and intellectual dis/abilities of various kinds. I would add that it might also include "physical" illnesses accompanied by mental effects (for example, the "brain fog" that attends many autoimmune diseases, chronic pain, and chronic fatigue)" (2011, 18).

not desire in the right way" or those who are alienated by their failures to adhere to emotion norms (ibid 240 n. 31; 189).

Affect aliens feature prominently, though perhaps not by name, in the moral philosophy literature, particularly in the works of those authors seeking to establish who is in and outside of our moral community. The scepter of the psychopath (or the sociopath, or the person with antisocial personality disorder — the minutiae seem not to matter) haunts introductory moral philosophy courses; because they fail to feel correctly, they fail to *act* or *be* morally correctly. There is another group, too, who often appears as affect aliens, though I want to caution that this might not be the best way of considering their affective agency: neurodivergent, and particularly autistic, people. Diagnostically, there are two central components of autism spectrum disorder(s) (ASD): "persistent deficits[37] in social communication and social interaction across multiple contexts" and "restricted, repetitive patterns of behavior, interests, or activities" (APA[38] 2013, 51). The former are of particular interest for this project and importantly include "deficits in social-emotional reciprocity" (ibid). Such deficits can and do pan out in a number of ways, but what is relevant here is that they can present as a difficulty in understanding social cues, norms, and expectations, including those dictated by the kinds of emotional meta-norms mentioned above. Social norms are a good example of reciprocal social communication, because they pertain to a person's ability to take in whatever communicative inputs they are receiving from another (e.g., of their emotions) and respond accordingly.

Before continuing, I should note that there is sometimes a conflation between autistic persons having difficulty understanding and acting in accordance with social cues and their

[37] This definition is not accepted by many autistic scholars and many others who write about disability including neurodivergence. The 'deficit model' suggests that autistic individuals are always fundamentally lacking in some capacity, rather than that they have (e.g.) some semblance of socioemotional lives, but that they might just enact these lives differently than neurotypical individuals. See Yergeau 2018; Kapp et al. 2013; Chapman 2019; Chapple et al. 2022).

[38] Here I am using the DSM's diagnostic criteria for autism spectrum disorder. This criterion is not without criticisms from the autistic community, and the reliance on medical diagnostic criteria is more broadly criticized by many within disability justice circles.

lacking empathy. That is, because autistic people have a different relationship to social norms including emotion norms (and often fail to mimic such norms in the same ways that neurotypical people do), they are sometimes taken as incapable of empathy. Aaltola (2014), Bollen (2023), and Tewes (2024) are among the authors who explore and critique this idea. Bollen calls it the “neurotypical gatekeeping of the concept of empathy” when empathy is understood in such a way that it fundamentally excludes autistic ways of being in the world (2023, 643). To say that autistic persons are not capable of empathy is a huge misstep, but it is one that follows naturally from overreliance on social norms as indicative of emotional-moral standing. Aaltola considers the question of autistic empathy in the context of authors like Kennett (2006), but instead of following his lead and concluding that autistic people are incapable of empathy, she concludes that such claims rely on an overly narrow conception of empathy. If we allow that there is a way of understanding empathy which involves not feeling as others feel (because they feel that way) but instead considering others are valuable in themselves, then autistic people can be and often are empathic (Aaltola 2014, 88).

Even though autistic persons are by no means precluded from empathy, the supposition that they are often follows on the heels of recognition that they experience social cues differently than non-autistic persons do. This ties in with the first diagnostic criterion mentioned above (“deficits in social communication and social interaction”) and often renders autistic people *affect aliens* relative to neurotypical (i.e., not autistic or otherwise neurodivergent) people. Responsive norms like those described in the previous section are a type of social norm because they dictate whether (and under what set of circumstances) a behavior is appropriate. The relevant behavior here has to do with responding to another’s emotions; for instance, whether it is appropriate to offer support and comfort, gauging whether a person wishes to talk about their emotions or be left alone, and so forth. If responsive norms are a species of social norm, and if social norms are often exclusionary (sometimes unintentionally so) to people with autism, we can conclude that many people with autism are excluded by responsive norms.

I do not want to suggest that these norms being opaque to many autistic persons makes them (the norms) *necessarily* unjust. Indeed, I think that moral prohibitions against (to use an extreme case) unprovoked murder apply even to individuals who do not or cannot recognize them. Instead, I want to point out that when such norms become expectations without any critical reflection, many people who cannot or do not accept them quite as easily might be marginalized or (to use Ahmed's terminology) alienated. Even worse, those who fail to easily adopt social norms can experience violence from caregivers or others (Yergeau 2018, 79-82; Larson 2021). To summarize: responsive norms dictate not what particular feeling is appropriate in a situation, but how individuals ought to behave in certain situations, particularly with respect to others' displays or testimonies of their emotions. Regardless of whether autistic persons feel their own emotions in the same ways that neurotypical people do (and this is a topic of frequent debate) it stands that they tend as a group to experience more obstacles in recognizing and adhering to social/behavioral norms, which include those norms dictating how to respond to others' emotions as described above.

Affective Injustice and Social Norms

The affective injustice literature as described in the previous chapter cannot account for responsive norms that govern our dealings with others' emotions. There are three ways of understanding the reason for this: first, because affective injustice is about individuals' own emotions and not how they meet the emotions of others; second, because the affective injustice paradigm at present is unable to make sense of non-uptake based injustices (i.e., those that do not require a sentiment's being expressed to the victim), and the norms described above can exist independently of any particular utterance and (thus) of the uptake of any particular utterance; and third, because social norms including responsive norms exist and are enacted even among equals because of their internalization.

To illustrate the first of these, it makes sense to begin by considering who the 'victim' of affective injustice is. Of course, this runs the risk of oversimplifying by rendering complex injustices dyadic (as relations of a single victim and a single perpetrator), but this oversimplification is worth it for the purposes of figuring out who fares poorly in each instance of affective injustice. The three paradigms that I will concern myself with here are Srinivasan's, Whitney's, and Gallegos', though I believe that any picture of affective injustice can be broken down similarly.

Under Srinivasan's understanding of affective injustice, the victim of affective injustice is the person who is forced into the normative conflict; the victim must choose between their emotions and bettering their circumstances. Because Whitney's affective injustice is grounded quite clearly in uptake — namely in uptake being given only in part or trepidatiously (see Whitney 2024)[39] or being denied altogether — the victim-perpetrator roles are clearly delineated. In most cases, the victim is the person whose emotion is not given proper uptake. There are some exceptions, of course; in the case of mis-metabolization, the victim might be the person who is forced to contend with someone else's excess emotion (Whitney 2018, 499). The perpetrator seems slightly clearer still here than it is in Srinivasan's picture; although I do think there can be a structural dimension to uptake denials (e.g., the ways that it is more difficult for some groups to be heard and believe than it is for others), many cases consist of an individual outright refusing to acknowledge (*take up*) someone's emotions. Or, in cases where uptake is not outright denied but someone is compelled to metabolize someone else's excess feeling, the person to whom the emotion initially 'belonged' would be the perpetrator. Srinivasan's picture of affective injustice does not require a clear-cut perpetrator,[40] as individuals can be compelled toward (e.g.) silencing their anger through broader social cues. For Gallegos, affective injustice lies in the

[39] "To be sure," she writes "there [are] limitations to a binary opposition between giving and refusing uptake" (2024, 46). The grace we ought to give others' emotions, even when those emotions can be read as (e.g.) violent or spiteful, varies depending on a wide range of social, political, and cultural circumstances. The anger of someone who experiences structural oppression thus deserves different kinds and degrees of uptake than a racist's anger.

[40] To be sure, there are cases where the person perpetuating the normative conflict (that is, the person instantiating it at a given moment by buying into it) but theories of structural injustice tell us that this person might not be strictly speaking *to blame* (see Young's *Responsibility for Justice*, 2011/1990).

deprivation of fundamental affective goods. In the cases he gives, the victim is the person who is denied affective goods. In Huxley's *Brave New World*, to use one of Gallegos' examples, the characters who lose the ability to experience certain emotions ("anger, fear, and sadness") are the victims of affective injustice because they are deprived of certain affective freedoms. Like in other paradigms, the perpetrator is not quite clearly defined here, even if the victim is more so.

Beyond what is discussed in the previous chapter, what do these pictures of affective injustice have in common? My contention is that they are all fundamentally about emotions simpliciter. Even in cases like the mis-metabolization mentioned above, it seems that for Whitney what makes this unjust is that those (to reiterate, victims) who are coerced into metabolizing others' emotional excesses is the emotional disturbance wrought by contending with another's emotions, particularly if those emotions are themselves grounded in injustice (like a Black person having to contend with a racist's *racist anger*). Even here, the at-issue phenomenon is an emotion simpliciter which, as mentioned above, is not the primary focus of responsive norms. That is, even though Whitney allows for cases where affective injustice arises because of someone else's emotions (rather than solely because of the treatment of the victim's own emotions), she seems to ground this in the targeted individual's emotions still. She dubs these the "unique *harms* of being *burdened* with affective waste" (Whitney 2018, 499; emphasis added). This is still affective in the simpler way because the language of being burdened is passive. There is no mention here of whether the burdened person is expected to do something with their burdens; the burdens themselves, qua affective states, are the relevant harm. As outlined in the previous two sections, therapy-speak — particularly carelessly applied RTS — is not about emotions simpliciter. It is, rather, about social norms that are adjacent to (but not, strictly speaking, *about* or governing) such emotions.

The second and broader reason that responsive norms are not explainable in terms of affective injustice is because the extant models of affective injustice are all *uptake-based*, either explicitly or indirectly, and so they locate affective injustice in particular (outward, expressive) acts. Responsive norms can be *related* to uptake, to be sure; oftentimes they dictate norms for how to provide uptake to others' emotions. However, they are not uptake-based in the same way; they do not require it. The victim of affective injustice — the person 'experiencing' the injustice – is the person whose emotions are taken up wrongly, either by being denied[41] uptake or by being given only a sort of truncated or warped uptake. Social norms — including those expressed through therapy-speak — are not about an emotion receiving or not receiving uptake, but about an emotion being recognized (so perhaps receiving some base level of uptake) and then prescribing what an individual is supposed to *do* or *say* after the fact of uptake. Focusing on uptake (or failures thereof) do not quite land us at the correct level of analysis, because we want to presuppose some level of uptake and look at what happens thereafter.

Normalizing Injustice?

This brings me to the crux of this section, and of the first half of this dissertation. My concern is that an attitude of deference toward the emotions tends to produce or compound harm, and thus to violate affective ethics, for two reasons. First, as mentioned above, given that conversations about empathy tend to rely on the emotions, and that the emotions are tangentially related to the responsive norms that we rely on when deciding how to respond to them, there seems to be a natural inference that those who do not adhere strictly to social norms that are *related to* emotions, even though not strictly *emotional*, are incapable of empathy. Being perceived as incapable of empathy can have serious social consequences, given that we think of empathy as a prerequisite for forming relationships and for membership in a moral community. If we rely too heavily on responsive norms (and adherence thereto), those who are unable to adhere to such norms are alienated. They are viewed as outside of the moral community because their *perceived* lack of empathy can give others the impression that they are incapable of recognizing their needs and desires.

[41] On my understanding, this includes the mis-metabolization discussed in the previous section.

Second, and perhaps more controversially, I worry that expressing responsive norms through platitudinous utterances like those enveloped by RTS is too uncritical of emotions' normative standing. There are some emotions to which RTS expressions are a fitting response, but I worry about creating some sort of imperative to validate or recognize as neutral *unjust* emotions. These might include those emotions described in the first chapter of this dissertation, those that exist within and uphold the ableist imaginary. If we become too accepting of the idea that everyone's feelings are *valid*, does this not force us to validate oppressive emotions? Must we concede that the ableist's pity, no matter how intensely and even violently it presents itself, is *valid*?

This is why I believe it is imperative to construe emotions as normatively evaluable. This does not mean that we need to think about them in terms of aptness (and thus adherence to emotion norms simpliciter) but that we should be able to claim that some emotions are more or less *just*. The opportunity to make such analyses has been overlooked by those who have already written about affective injustice. Emotions can cause harm, as can the norms we let dictate our responses to them. Attending to these harms requires that we lend credence to the ways that emotions can themselves be *unjust*. This does not require us to cede that emotions are fully under individuals' control; they are not. Instead, it requires us to see emotional tendencies as malleable, as trainable, so that we can come to feel more justly.

Conclusion

In this chapter, I have noted two distinct categories of emotional norms: those governing emotions themselves, and those that govern the nexus of behaviors surrounding emotions. Much work in feminist philosophy has focused on the former, including by authors who note racist and sexist prohibitions against particular emotions like anger. Moreover, much has been written about the ways that emotions become racialized and (thus) deemed apt for some racial groups yet not for others. A white person might be allowed to be simply angry, while the anger of a person of color is consistently transformed into rage — which carries more heft — or Madness, itself rife with normative implications because of stigma against mental illnesses. Whether a person is allowed to feel (and express) anger depends not only on what has happened, but on who they are.

The second arm has received considerably less uptake, likely because it appears as a sort of social norm simpliciter rather than being contained in discourses about the emotions. These norms are nonetheless worth paying attention to beyond the attention already paid to social norms, though, because of the implicit ways they are passed on. There is often not any explicit statement of the norms, but there can be serious social consequences for those who fail to meet them. This is particularly harmful when we consider groups that have tenuous relationships to social norms, as these norms' proximity to emotions — even though, again, they are not strictly speaking emotional norms — and the relationship between emotions and empathy mean that those who fail to adhere to responsive norms are often perceived as not possessing (or being capable of possessing) empathy. Given that empathy is often used as a metric for moral agency, individuals who fail to adhere to responsive norms are often excluded from participation in moral communities that they should be readily accepted into.

In what follows, I take two further steps to expand upon the arguments I have made thus far. First, I explore the digital mediation of emotions, and the ways that online spaces encourage users to be apathetic. Then, in the fifth and final chapter, I begin making a positive argument, suggesting *attentive care* as an ameliorative affective disposition that clarifies the obligations we do have to others' emotions without reifying unjust hierarchies in the ways that RTS does.

Chapter 4 - (Dis)Content Creators

Introduction

The three previous chapters have identified a set of phenomena that, though they span a wide range of cases, can all be classified as showing the necessity for an affective ethics, or an account of how we ought to treat others with respect to their emotions. Taken as a whole, these three chapters explore (respectively) the ways that emotions can be the substance of wrongdoing, extant accounts' attempts to systematize the ways that we can do harm with respect to others' emotions via the affective injustice framework, and the ways that such harm extends to cultural scripts surrounding (responding to) the emotions even when they are not, strictly speaking, emotion*al*.

My target of analysis in this chapter is the circumstances under or environment within which such wrongs can proliferate or are made more likely. I am concerned especially with online environments, and how their features – particularly when viewed in contrast to 'real life' ones – makes them the site of (affectively) bad behavior. I work from an assumption that many emotion-related wrongs — including and perhaps especially those described in the three prior chapters — can be understood as failures to fulfill a duty to care, and that they can be understood as expressing apathy. The exact contours of this duty to care (and how we might use it to motivate an affective ethics of attentive care) are studied in the subsequent and final chapter of this dissertation. For the purposes of this chapter, suffice it to say that we should care about others, and that some level of care toward others' emotions is laudatory if not obligatory, even if we have not yet sketched out precisely what we mean by 'care'. Moreover, I assume throughout this chapter that care and apathy are at odds with one another; to care for or about someone or something is to have some regard for them or it, and the lack of such regard means we are apathetic.

What is at issue in this chapter, then, is the extent to which cyberspace (or online space) makes fulfilling this moral obligation to care considerably more difficult. I focus on social media platforms, as they are among the areas where users spend much of their (our) online time. In what follows, I identify three structural levels on which social media platforms disincentivize care and (thus) incentivize apathy, and other dispositions antithetical to care including resentment or hatred. These three levels are structural in the sense that they refer to the infra*structure*[42] of social media, the ways that platforms are organized and operated (that is, the ways they are actually *structured*). These structures are often taken for granted (in other words, we assume that they represent how social media simply *is*) but they quietly shape the ways we are accustomed to caring (or not caring) within the spaces they delineate. They set our expectations within those spaces. The three levels that I identify get increasingly granular, beginning from the decision to use one platform rather than alternatives and ending with the content that users come across on a platform.

First, and most broadly, I demonstrate that **patterns of social media use** discourage care, arguing that users' tendencies to self-select into media repertoires leads to conceptual fragmentation because of a lack of cross-platform interoperability. That is, individuals might be pulled toward apathy toward those they do not interact with because of their self-selecting onto different (groups of) platforms. Second, the incentive structures of social media operate *between and among users* to facilitate **the creation of weak, rather than strong, ties,** which necessitate less sustained relations with others. Accordingly, because we encounter online others as abstract rather than concrete, it is more difficult to sustain care toward them, even if we construe care as having minimal actionable commitments. Third and finally, I argue that

[42] Rather than being structural in the more normative sense in which philosophers tend to use the word. These two uses of 'structural' are not unrelated, but I mean it in the more literal sense.

(likely owing to the first two broader structural elements) the **content** that's shared on social media tends to be affectively loaded, but platforms themselves provide inapt spaces for mediating such content, and so responses default to apathy rather than engagement. In other words, the three structural levels on which individuals encounter apathy on social media can be boiled down to: individual relationships *to* social media, individual relationship dynamics *on* social media, and the content that individuals produce *through* social media. I do not mean to suggest that, at least on their own, any of these practices are problematic, or that any of them lead unfailingly to carelessness. Instead, the three structural dimensions of social media operate in tandem to create a superstructure that encourages apathy rather than care. I have deliberately chosen these three structural elements of social media in part because they inform one another. That is, media use patterns determine relationships on social media because they determine who the other users (who become party to those relationships) are in the first place, and who is on a platform plays a determinative role in what content proliferates on that platform and how (or to whom) it spreads. Likewise, platforms can gain reputations for hosting certain types of content — especially when it is content that might be prohibited by other platforms' terms of use — which can reinforce media-usage patterns.

This chapter is motivated in large part by a line of inquiry posed by Sherry Turkle in 1996, concerning the supposed tension between virtuality and real life. She writes that: "To the question, 'Why must virtuality and real life compete—why can't we have both?' the answer is of course that we will have both. The more important question," Turkle adds, "is 'How can we get the best of both?'" (1996, 53). This chapter thus serves as a means of asking how the ways that our social media are currently structured gets in the way of our reaping the best that both virtuality and real life have to offer.

Social Media, Content Creation, and Incentives

What is social media?

Most people, when faced with the question "what is social media?" will immediately point to popular platforms, so ubiquitous that definition seems to be redundant. 'Social media? Oh, Facebook.' Facebook, Twitter/X, and Instagram are undoubtedly social media platforms, as are newer-age decentralized platforms like Bluesky and Mastodon and platforms of yore like Myspace and Xanga. Platforms like Reddit and YouTube probably are too. Definition through exemplars can be difficult, though, because it fails to convey a "consensus on what *defines* these tools as social media," and so makes it difficult to determine whether a new platform meets the relevant criteria and "to develop broad, robust theories" about social media usage rather than a particular platform's usage (Carr and Hayes 2015, 46; 47). In this relatively brief section, I raise some general criteria for what social media is, drawing from other theorists' past attempts at defining the term. I do so in part owing to Carr and Hayes's contention that "the lack of a common definition can result in multiple connotations of a concept, making it difficult to create a shared understanding [Hempel, 1966] to guide theory and research" (2015, 48). In other words, it is futile to theorize about things that happen *on* social media without at least speculating about what social media, target of analysis, actually is. In short, I understand social media as the set of online (taken to encompass both traditional web-based and mobile/cellular connectivity) platforms wherein users' primary means of participating is through content-creation (often in the form of 'posts') connected with a relatively consistent individualized user profile. Below, I explain the origins of such a definition.

First a brief note on terminology: there is disagreement within the literature about whether *social media* and *social network(ing sites)* are synonymous, and about how exactly these differ from earlier technologies like blogging sites (Kane et al. 2014, 276). For those who treat the two as distinct, the former is typically treated as the broader category encompassing the latter; on this view, social networking sites are a *type* of social media, and social media might include web functions like e-mail. For my purposes, though, I will treat the two terms as more or less equivalent; when I use the phrase *social media,* I am merely using it as a shorthand for *social networking sites/platforms/applications*.

One definition comes from boyd and Ellison, who "define social network sites[43] as web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system" (2007, 211). To think about these requirements in terms of an exemplar, consider Facebook. The site is undoubtedly web-based, emerging from Web 2.0 applications' "shift from user as consumer to user as participant" (Obar & Wildman 2015, 746). On Facebook, users construct profiles containing different artifacts of personal information: name, photo(s), a brief biography, location, education and work histories, language(s) spoken, and so forth. By using the "Add Friend" feature, they articulate their connections with other users. Using the "Like Page" and "Join Group" features allow them to do this in slightly different ways, joining communities based on (e.g.) brand loyalty, shared ideas, or common interests without necessarily 'friend-ing' the other group members. Finally, the site's interface allows users to "traverse" their connections, visiting individual profiles/pages/groups. Moreover, individuals can see their friends' friends and can see the other members of groups they have joined. Though Facebook has changed quite significantly[44] in the two decades since its founding, these core features have remained relatively stable.

Boyd and Ellison note that their definition allows for variation across social network sites; for instance, not all sites require bidirectional Friend-ing. That is, though Facebook requires both individuals to consent to becoming friends (one by initiating the friend request and the

[43] The earliest social network site *proper* that the pair identifies is SixDegrees.com, launched in 1997 (2007, 214).
[44] For instance, through the addition of sponsored posts *to* users' news feeds rather than promoted content being limited to smaller ads in a side bar.

other by accepting), other sites like Twitter/X do not, particularly for public (that is, unprotected) accounts. Carr and Hayes attempt to define social media as a broader category, defining it as "Internet-based channels that allow users to opportunistically interact and selectively self-present, either in real-time or asynchronously, with both broad and narrow audiences who derive value from user-generated content and the perception of interaction with others (2015, 50). There is quite a bit of overlap between this definition and the one offered by boyd and Ellison—both posit that their targets are web-based, that they involve user profiles (self-presentation can be understood as constructing a self-profile), and that interaction and connection are among the foremost goals of such sites. Carr and Hayes make explicit a temporality of social media, though I do not think that their addition is something that boyd and Ellison would have rejected in their earlier definition, given that many of the examples they describe are similarly equipped for asynchronous interaction. Though Carr and Hayes bemoan the imprecision that conflating social network sites (boyd and Ellison's target) with their own social media creates, they fail to set forth precisely what theoretical disservice we do by treating the two as coextensive. Obar and Wildman are explicit in their treating social media and social network as equivalent. They define social media—their and my term of choice—as having four defining features: (1) They are Web 2.0/Internet-based; (2) "User-generated content is the lifeblood of social media," that is, is the reason individuals return to platforms; (3) Individuals and groups establish themselves on social media by creating profiles; and (4) Social media creates online social networks by allowing individual and group profiles to connect and interact with one another (2015, 746-7).

(Why) do social media and its incentive structures matter?

With some understanding of what social media is, we come to a perhaps weightier question: why does social media *matter* as a target of academic (and especially philosophical, rather than technical) analysis? Part of my answer to this is fairly straightforward: social media matters, particularly for a project focused on harms that arise with respect to the emotions, because (1) many wrongdoings occur or are processed in real-time on social media platforms and (2) even outside of these contexts, such platforms comprise a not-insignificant part of many persons' lives, particularly those in younger generations.[45] The importance of social media should not be understated, then, given that it is where so much of our lives *happens*. Megan Fritts writes that "as of 2017, data surpassed oil as the most valuable asset in the world" (2023, 14). That users are so active on social media platforms means that they are generating massive, previously unthinkable amounts of data for the tech companies heading such platforms. As will be discussed in more depth below, this data is used to make algorithmic determinations about what content a user is presented with (or, as Gillespie [2020; 2022] shows, which content features less frequently on their feed), which itself shapes — reinforces, even — individual media-use patterns. It keeps users coming back.

Social media is also remarkable because of the scale at which it brings users into contact with one another; communities are not limited by geographic location, or by the slower timescale of (e.g.) the postal system as they once were. Moreover, the immediacy of social media — that it allows for near-real-time engagement among users, including about breaking news events — means that expressions on platforms are might not be as cultivated as those on other forms of media (e.g., print journalism, where pieces are often produced on a time delay and with at least one level of editing/checking). In this sense, social media might give us less mitigated access to expression, and its content might thus express knee-jerk, intuitive emotional responses to events. As made clear in the preceding chapters, I am especially intrigued by individuals' responses to injustices, how they represent individuals' internalization of harmful depictions of marginalized groups.

Lastly, particularly in the past decade, social media has taken on a new role in public life. I would be remiss not to mention the explosion of mis- and dis-information on online platforms

[45] 2023 Gallup research found that US teenagers spent an average of 4.8 hours per day on social media. A majority of adults under age 30 reported daily social media usage (Gottfried & Park 2025).

in recent years. Eriksson Krutrök and Lindgren point out that the mediatization of contemporary life - the fact that so much of *what we do* is done *online* — makes "information holes" especially dangerous, because such holes can be exploited by bad actors looking to spread falsities (2022, 81-2). The mis/disinformation that comes to populate such lacunae then spreads rapidly because of the ways that social media platforms incentivize engagement in a quantitative, rather than quality-first, sense; it is often considered better (including for monetization) to go viral than it is to garner a mere handful of thoughtful responses, and provocative content is more likely to go viral than is more measured content. One of the most prominent examples of this — and one of the earliest at such scale — occurred in the lead-up to the 2016 US presidential election, after which Facebook came under fire for allowing "fake news articles" which "users thought…were real" and thus shared widely with their social networks to influence the election in favor of the eventual victor, Donald Trump (Fritts and Cabrera 2022, 456). While a full-length analysis of mis- and dis-information is beyond the scope of this chapter, the idea that falsities spread on social media platforms can influence elections for one of the world's most influential positions makes it clear that social media is no longer merely something that people do alongside their 'real lives'; it is enormously consequential for our democracies (see Brown 2018), and for all that hinges upon them.

What does it mean to say that social media platforms *incentivize* apathy, or that they *disincentivize* care, defined in more depth in the next chapter but generally understood as good-faith, thoughtful engagement with others in a way that balances recognizing reality with respecting others' wishes for how they are viewed? Incentive structures are essentially choices made (often by platform designers) to nudge users toward one possible action rather than another. They make one option more desirable, either because it is easier to attain, or because users are rewarded in some sense for choosing it. At least a minimal form of such paternalistic — and I use this term descriptively and not normatively — interference is unavoidable (see Thaler and Sunstein 2003). Some social media incentive structures are relatively straightforward —

likes, reposts, follower counts — but may not accord with users' initial intentions upon joining a platform (see Nguyen 2024). Others are more difficult to measure, like the comparable quality of different relationships, and the precise ways (and quality of the ways) that we perceive other users. Moreover, incentivization is not plainly deterministic. That is, a particular course of action being made easier or beneficial does not (at least to an extent) mean that it becomes the *only* possible course of action. This is important to maintain for this chapter because it allows us to maintain that individuals do have the power to act differently on social media platforms even if they are nudged toward acting a certain way; otherwise, these choices and actions would be immune to moral critique.

(Social) Media Repertoires: Patterns of Individual Use

The notion of a media repertoire was first proposed by Uwe Hasebrink and Jutta Popp as a means of generating a more collectivist approach to studying individuals' media use (2006). That is, rather than look at uses of a particular media platform (television, print news, news radio, internet news, etc.) and the behaviors associated with it, the authors study "how people combine contacts with different media and different kinds of content," motivated by the fact that most users do not restrict themselves (or, more specifically, their media use) to a single platform (Hasebrink & Popp 2006, 369; Webster 2014, 37). For instance, given that many households that have cable television also have access to the internet, an avid watcher of cable news might also be a zealous consumer of online content. A singular media repertoire, that is, an *individual person's* media repertoire, "consists of the entirety of media he or she regularly uses," so repertoires differ among individuals (Hasebrink and Domeyer 2012, 758). Repertoires are, furthermore, shaped by the frequency of use of specific media, when and how decisions to use one form of media rather than another are made, and users' normative evaluations of different media. For instance, users' repertoires more prominently represent media they believe to be truthful, or on which they more readily find community, and so they come to be value-representative. The table below (Fig. 1) provides an overview of the components that factor into individuals' decisions to add a particular media form to their repertoire.

Table 1: Analytical framework for the description of media repertoires

Main category	Subcategories
Relevant components of media repertoires	Media types | genres | topics | concrete products/brands | social contexts
Empirical indicators for the description of repertoires	Contacts | preferences/attitudes | embeddedness in everyday routines | habits | 'addictive' behaviour
Relations between the components of the repertoire	Proportions of use | relevance | diversity | functional complementarity | compatibility vs. competitiveness
General principles guiding the composition of media repertoires	Selective vs. unselective | hedonism vs. conscientiousness | high vs. low prestige | decisions not to use certain media
Overall subjective sense of the media repertoire	Subjective theories | social context | personal values and ambitions

Fig. 1 (Hasebrink and Domeyer 2012, 760).

Media repertoire studies explore the "bivariate correlations between the amount of use of two media," or the likelihood of an individual who makes frequent use of one media also making frequent use of another (Hasebrink and Popp 2006, 374). They thus expose "comprehensive patterns of media use," because the media repertoire approach creates clusters of media (and thus of users) based on those which have higher likelihoods of concurrence[46] (ibid). Media repertoires as a research lens also draw attention to the relationship between demographics and media use; whether, for instance, people of a particular age are more or less likely than others to consume particular sorts of media (cf. ibid 378). Matassi et al. find, for example, that women are more likely than men to make use of social media platforms (2022). Women will then be more likely to include in their repertoires media that are regularly accessed *through* social media

[46] For instance, there might be relatively little overlap between Fox News and CNN viewers, but considerably more between CNN viewers and *New York Times* readers.

platforms (through, for instance, click-throughs that make content from one platform accessible from another).

Media repertoires are meaningful in a deeper sense, too, insofar as individuals construct their social identities via the media they make use of; if given the message that a certain type of person reads *The New Yorker*, a *New Yorker* reader may come to find themselves further approximating this archetype. Luthar and Črnič expand upon the final row of the above table (subjective sense of the media repertoire) arguing that media repertoires or — as they call them — media *ensembles* "relate to" and express "various cultural competencies and cultural codes, as well as ideology, that is, political views and values" (2017, 417). The authors focus in part on the social dimensions of media (including but not limited to social media), drawing attention to "users' production of content *for other users*" as a "social act whereby viewers, readers and users are constituted in relation to media," rather than being mere receptacles unto which the mediated information is projected (ibid, 420). Thus, media repertoires are necessarily value-sensitive, because individuals generally add to their repertoires those platforms on which they see their own values and views reflected. More than simply recognizing that media repertoires exist, I want to explore the ways that they might fragment individuals in such a way that discourages care or encourages apathy. This does not, to be sure, mean that there is anything *wrong* with repertoires themselves; the apathy I describe is simply one of the downsides of the identity cohesion they facilitate. There are upsides too, including the possibility of users finding like-minded community members. This research requires focusing on the social dimensions of media repertoires as expressed by Luthar and Črnič. As evidenced by the framing of this chapter, I am primarily interested in *social* media repertoires, understood as the patterns of social media/social networking platforms that individuals use.

Social media repertoires as a distinct concept is relatively novel, at least by name, though social media researchers have studied users' decisions to use and share certain parts of their lives on some platforms rather than others; some content is simply better suited for

Instagram than for Facebook, for instance (see Matassi et al. 2022 for a survey of such research). Such repertoires are nonetheless worth studying as repertoires because people use social media "in a more stable pattern, a product of habit or routine formation over time," likely owing to the ways that such platforms are designed to drive repeated engagement[47] (Matassi et al. 2022, 134). Media like newspapers and television are *de facto* excluded from consideration in this project, then, while the category of social media – which more traditional repertoire studies may consider as a cohesive unit – is broken up to include different platforms: Facebook, Instagram, Twitter/X, Reddit, TikTok, and so forth. In the last several years, such research has also begun to consider decentralized platforms like Bluesky and Mastodon. Each individual social media platform counts as its own medium, and so even a person who *exclusively* consumed media via social media platforms could have a quite expansive[48] and diversified repertoire. The latter point follows the example set by Webster (2014), who resists the urge to view broad swaths of media – television, for example – as repertoire 'units,' instead drawing attention to the ways that individuals' consumption even within a single media type is idiosyncratic, as (e.g.) frequent consumers of television news might be situated quite differently depending on whether they primarily watch Fox News or MSNBC. What follows in this section might speak of *media repertoires* rather than the narrower category of *social media repertoires*, but I want readers to keep in mind that my target is the latter, and so when I refer to repertoires more broadly from this point forward, I am using the phrase as shorthand for social media repertoires. The former simply figures more in the literature because of social media's relative novelty.

That there are demographic differences across social media repertoires – that, for instance, women of a particular age range might tend toward one platform rather than another

[47] I assume that there is also something of a sunk cost fallacy at play here; because individuals may have already dedicated ample time on one platform to building a presence and community, they may have little or no desire to begin 'from scratch' elsewhere.

[48] Further research might explore the simultaneous expansion of social media repertoires alongside centralization and interoperability; the notion that users might be joining a greater number of platforms, but that those platforms are quite closely integrated with one another.

and thus have it in their repertoires at a higher frequency, or that platforms tend not to be populated by the world's most impoverished classes because of disparities in broadband access (see Sherry Turkle's early warnings about the creation of an "information elite") – means that individual platforms *de facto* create spaces wherein users have at least *something*, however general and even tautological, in common (Turkle 1996, 57). This is reinforced by Kane et al.'s finding that users of the same platforms tend to behave similarly, though they attribute to this not to self-selection onto the platform itself but to users' exposure to the same content after having chosen which platforms will construct their personal social media repertoires (Kane et al. 2014, 282; Borgatti and Foster 2003). Luthar and Črnič find a correspondence between media repertoires and political attitudes showing that, for instance, members of the predominantly upper-class "media integrators" repertoire – defined as consisting of "those who are deeply involved in both digital and analog media culture" – hold more consistently liberal positions regarding "the civil rights of the LGBT population, women, ethnic minorities or migration" than those who remain squarely within either digital or analog media (2017, 425; 432). If social media platforms come to be 'nodes' of contemporary media repertoires, there is reason to think this pattern is replicated in the relatively new media; hence the (well-founded) charge of platforms enabling *epistemic bubbles* and *echo chambers*[49] in which individuals' views are reinforced through exposure primarily to complementary ones, often paired with reason to distrust conflicting ones (Nguyen 2020).

I do not want to argue that it is necessarily bad for certain populations to tend toward certain platforms, particularly given the necessity of demographic-specific platforms[50] for reasons of solidarity and consciousness-raising, as well as for providing spaces where group

[49] These phenomena complement one another. Nguyen defines epistemic bubbles as "social epistemic structure[s] in which some relevant voices have been excluded through omission," and he defines echo chambers as "social epistemic structure[s] in which other relevant voices have been actively discredited" (2020, 142).

[50] Consider Rudy Fraser's Blacksky Algorithms, which "lets communities create safe spaces on their terms" (blackskyweb.xyz). Fraser initially founded Blacksky on the same AT Protocol as emerging social media platform Bluesky, with an eye toward creating a safe and self-governed space for Black users.

members are insulated from certain violent or hateful speech. However, we should acknowledge that this might amplify negative attitudes toward out-groups (understood not normatively but as those without a particular platform in their media repertoires) through the creation and reinforcement of the in-group. That is, if certain demographic groups tend to use certain platforms, then some users of those same platforms are *de facto* (and, I add, in a probably morally neutral way) rendered out-group members on those same platforms. It is worth noting here that I use the terms in- and out-groups in ways less value-laden than their traditional uses. Normally, in-group refers to a privileged group, while the outgroup is those who are marginalized. However, I am taking the 'in' and 'out' more descriptively and less normatively. The in-group relative to a platform is simply the users of that platform, while the out-group is those who do not use it at all, or who use it so sparingly as to render their uses of it negligible; perhaps a person who made an account then deleted the app altogether. The out-group in this sense has not *necessarily* had some wrong — such as the wrong of marginalization — done to them, because their exclusion is often (though not always; consider age-restricted access or regional internet and platform shutdowns/blockages[51]) technically self-imposed in that they elect not to join a platform.

Though the creation of social groups as such is a natural consequence of public spaces, including those in the digital public, and thus should not be discounted altogether, group 'clustering' on certain platforms may cause out-group members who join – those who do not fit a platform's primary demographic(s), even if the platform was not created *for* a certain demographic – to be alienated. This is the problem with epistemic bubbles as Nguyen sees them, though his concern seems to be more with the ways this happens within established platforms (the self-selection of friends on Facebook, the platform where he receives most of his news), and

[51] The Open Observatory of Network Interference (OONI) tracks intentional internet censorship. https://ooni.org

not the more fundamental way that self-selection of platforms helps to reinforce this process (the decision to sign up for a Facebook account in the first place) (2020, 143).

How, though, might this sort of self-selection into media repertoires weigh on expressions of empathy (or the dearth thereof) on platforms? Guidance from the social sciences is helpful here. For example, studying the relationship between outgroup exposure and bias, Schieferdecker and Wessler (2017) find a general disinterest in outgroup-centered media among South African youth (1006). Though their study tracks the attitudes of a relatively small group (n=1,095), we might still take the authors' findings as evidence of the persistence of group biases—biases that, by consisting primarily of disinterest, can be characterized as apathetic—even in the face of widely available outgroup media. Moreover, if the relative homogeneity of a platform leads to a false assumption of consensus (as Luzsa and Mayr [2019] propose), then there may be greater levels of apathy toward those who do not fall into the (platform-specific) in-group, even if this group is constructed around something relatively tenuous. This is supported by the broader literature on empathy and apathy within and among social groups. Although "superordinate categorization" (categorization into broadly-defined groups) can lead to intergroup prosocial (that is, helping or caring) behaviors among smaller groups that share an overarching identity, if groups are especially strictly separated (i.e., if this superordinate categorization is missing or is not recognized), such behaviors might be discouraged or at least not encouraged, given that generally "empathy is reduced for outgroup versus ingroup members" (Thomas et al. 2009, 311; Lockwood et al. 2017, 1). Cikara et al. (2011) study the latter phenomenon in great depth, highlighting research on the limitations of empathy when studying intergroup interactions. When faced with "socially distant" others – those who we view as decidedly not like us, or as less like us than some *other* others – we tend to misread or altogether *miss* their emotions, thus truncating empathy (Cikara et al. 2011, 149). This is often colloquially referred to as an "us vs. them" phenomenon.

If "outgroup members…reliably elicit diminished perceptions of suffering" because we do not recognize their suffering as such and they therefore elicit less empathy than members of our 'own' groups, then we might assume that they elicit *apathetic* responses (ibid, 150). If use-patterns of social media platforms are such that they create in- and out-groups via the tendency for some groups to use some platforms without those platforms being wholly insulated (that is, unavailable to the so-called out-group), then the very existence of social media repertoires can encourage apathy. Accordingly, it becomes more difficult for individuals on such platforms to act caringly because the platforms are structured so as to discourage a caring disposition (Tronto 1993, 105). Put differently, self-selection into in- and out-groups via the creation of media repertoires might increase the *social distance* among groups (and thus among individuals in groups) (ibid, 37). Social distance, understood in the sense that Tronto uses it, refers in part to the proximity we view between ourselves and others qua moral beings. If groups perceive an insurmountable distance between themselves, then they will be less likely to recognize members of other groups as having the same kind of moral standing that members of their own group do.

Weak Ties: Relationships on Social Media

In addition to group-level patterns of platform selection via self-sorting into media repertoires leading to a normalization (perhaps even an *optimization*) of apathy, the kinds of inter*personal* interactions encouraged by the nature of social media platforms mimics the same pattern. This is no doubt related to the previous section's content, as the relationships that a person forms and maintains on a social media platform depends on who is available — who is using that platform — in the first place. Once individuals have self-selected into their platform repertoires, the relationships they find on those platforms (including even passing relationships with other users whose content they might come across) can, too, tend to cultivate apathy.

In short, social media platforms cultivate particular *kinds* of interpersonal relationships. For example, Yochai Benkler notes that the "new channels of communication" made possible by the networked information economy (NIE) can help to *thicken*, or reinforce, "connections with family and friends" (2006, 364). These are strong ties. Weak ties, by contrast, might consist of a series of one-off interactions without a requisite offline relationship, or without the online communications being substantive. In other words, where a relationship already exists, digital communication – emailing, instant messaging, and social networking – can make that existing relationship stronger. Benkler evinces this, noting "that people who were connected recognized three times as many of their neighbors by name and regularly talked with twice as many as those who were not wired" (2006, 363). This is supported, too, by boyd and Ellison's contention that "most [social networking sites] primarily support pre-existing social relations" (2007, 221). What is important, though, is that boyd and Ellison and Benkler are all talking about relationships that pre-date social media connection.

It is worth noting that these claims were made nearly two decades ago. In the context of social media's lifespan, this is quite the distant past. The prominence of influencer culture in recent years, hypervisibility, and the ways that 'going viral' (and thus having your online content exposed to people you do not have any pre-existing relationship to) has become a badge of honor mark how social media is (maybe paradoxically) shifting away from the social. The relationships forged on social media platforms are becoming increasingly *para*social and asymmetrical, not only with celebrities and public figures (see Archer and Robb 2025) but also with social media users who amass a sizable audience or following.

Though social media can certainly strengthen relationships—as with distant family members whom a user might rarely see in-person—and can forge entirely new relationships, it can alienate other individuals, not merely excluding them (which I do not see as necessarily wrongful; insisting otherwise would create an unjust expectation of access) but rendering them objects undeserving of empathy. I am chiefly concerned in this section not with cases like those of "virtual rape"[52] that Turkle (1996) describes, which create outwardly violent or hostile ways of

[52] In this case, Turkle is describing MUD users forcing sex and sex acts upon other users' characters/avatars, and to the use of sexually aggressive and violent language on such platforms (1996, 55).

relating, but with the more mundane-seeming antiegalitarian ways that cyberspace encourages users to interact with one another, and the ways that such alienating ways of relating complicate sympathy on social media platforms. Parasociality as mentioned in the previous paragraph is a case of this that I am increasingly concerned with, because it can be read as a sort of objectifying devotion to a user without a 'real' relationship. Archer and Robb define parasocial relationships as asymmetrical (in terms of attention, communication, and epistemology) personal relationships arising 'between'—though even this word seems unfitting—a fan or media consumer and a celebrity or media figure. Even without the kind of devotion that archetypal parasocial relationships entail, the specter of access—the ways that social media makes the people we are interacting with seem far closer (socially) to us than they are—means that online interactions are often colored with asymmetry because relationship expectations are not often clearly communicated, or they are communicated in situations where context, body language, tone, and so forth are far harder to recognize.

Why do relationships like this seem especially pertinent on social media platforms? I believe, following the lead of Norlock (2017), that we have yet to reckon with the "new responsibilities" that the internet introduces, including recognizing that membership in an audience grants users some soft power (188). Norlock takes up the notion of imaginal relationships – "relationships that window with imaginative content which includes their import, meaning, and membership" – to argue that, because cyberspace increases the base *number* of relationships that users are party to, there are increased opportunities for misbehavior, including online shaming (2017, 189). This of course is not to disparage well-deserved or righteous online shaming, which often befalls public figures (or even merely users with publicly available profiles) expressing hatred toward marginalized groups. Paired with the assumption that online harms are not real harms (a belief which seems to be held also by some of those Turkle describes), the expanded domain of interpersonality online makes it difficult to assess the harms that imaginal relationships (particularly those among the shaming group) can bring about.

Norlock notes that "the sense of distance from victims or uncertainty regarding the indefinite nature of targets," that is, a less-than-clear sense of *who* precisely the targets of our online speech are, "may ameliorate our individual feelings of responsibility" toward those persons (ibid). This uncertainty is exacerbated by the existence of "alt" or anonymized accounts where virtually everything about the user's identity can remain shrouded, at least to other users without backend access (see Lewis 2020). In these cases, we literally do not know who we are talking to besides knowing a moniker. We intuitively fail to view harms toward other online users as equivalent (as in, as equally harmful or wrong) to real-life ('IRL') harms. This is further exacerbated by the fact that online relationships (or relations, contact without an established relationship base) are a 'mixed bag,' so to speak. They are mixed in the sense of consisting of "partly voluntary and partly involuntary interactions," because even if a user's following is their primary audience, if an account is not set to private, the general public can still see (and thus engage with) content that is posted (ibid, 193). Phillips and Milner (2017) discuss a form of involuntary engagement, too, writing that the purportedly neutral "act of commenting on a story…risks amplifying the story's reach to other members of one's social network" (54). Commenting can be 'neutral' in the sense that it might express neither support nor disdain for a post's content yet allows for the memetic spread of whatever that content may be.

Involuntary engagement underscores Benkler's comments about thick versus thin (or strong versus weak) relationships; while he maintains, as mentioned above, that social media (via the networked information economy) can strengthen pre-existing relationships, he notes that it also increases the sheer *number* of ties or relationships that a user is party to. Many (most, even) of these ties, by virtue of psychosocial and even chronological constraints (I.e., that we do not have the time to forge strong ties to everyone we interact with online), will be weak. That social media forges a great number of ties, but that those ties might be relatively weak, points to a phenomenon that social scientists dub (among other terms) the media-empathy paradox. Guan et al. define the media-empathy paradox as “the irony that a tool created for social connection may be reducing connective capacities” (2019, 148). In other words, the *social* seems to have fallen away from social media in practice. Another way of explaining this same paradox is by considering two concepts: the contact hypothesis and compassion fade. The contact hypothesis is, generally speaking, the idea that coming into contact with groups other than your own (as we do on social media, particularly platforms where you’re exposed to content from accounts you have not actively elected to follow) can make you more empathetic toward those groups (Roberts 2021, 824). The contact hypothesis loses its explanatory power depending on scale via a separate phenomenon known as compassion fade. Compassion fade points to “the tendency for societal responses to decrease when confronted with large-scale humanitarian or environmental emergencies” (Thomas et al. 2018, 3782). In other words, as the number of lives in danger increases, compassion (which can be represented in part by the perceived value of any life-saving measure) tends to decrease (see Västfjäll et al. 2014). If we view compassion fade as a requisite fade in empathy or care for those others whose hardships we are exposed to, we reach the media-empathy paradox: even if social media *can* increase empathy by (e.g.) increasing awareness of various social groups’ circumstances, there exists a tipping point at which exposure actually lessens empathy.

Because social media allows for connection across physical and ideological distance, individuals’ social networks are no longer limited to those in their immediate family or community. Some – many, even – of these relationships simply would not be possible in pre-social-media times, or even prior to the advent of simpler digital communications technologies like e-mail because the latter does not involve the sort of randomized or involuntary contact that social media does. Even though you could, in theory, make up an email address that likely belongs to *someone* and send a message, the likelihood of response seems low, and this is not how the technology is primarily used nor is it how its creators intended it to be used. Benkler notes (optimistically) that:

> "[u]nless Internet connections actually displace direct, unmediated, human contact, there is no basis to think that using the Internet will lead to a decline in those nourishing connections we need psychologically, or in the useful connections we make socially, that are based on direct human contact with friends, family, and neighbors" (2006, 362).

What seems pressing, two decades after *The Wealth of Networks'* publication, is to ask ourselves whether the hypothetical in this statement has been realized; whether, in fact, online connections *have* come to replace in-person, analog ones. It seems that, even if they do not replace in-person connections particularly well, online connections have come to rival them quantitatively, especially following so many modes of communication shifting to virtual formats amid the COVID-19 pandemic (Marinucci et al. 2022). Thus, Benkler's optimism may not withstand the test of time, and there may be genuine cause for concern about a decline in positive ('nourishing') connections now, even if there was not in 2006. Even if platforms are not *intended* to supplant the in-person – though I do think there is reason to think that contemporary ones are – they have come to occupy this role.

Benkler cites Turkle's concern that "people engage in limited-purpose, low-intensity relationships" in online spaces (2006, 359). That is, even if the *quantity* of relationships increases owing to social media platforms' allowing users to amass hundreds, even thousands of 'friends,' this does not necessarily correspond to an increase in relationship *quality*. Consider the following possible relationship options between two persons (A and B):

1. A and B are more or less unaware of each other's existence. They exist and engage in totally separate spheres both online and IRL.
2. A and B have a friendly collegial relationship IRL, and follow one another on at least one social media platform, where they engage minimally, but positively, with each other's postings; they might occasionally 'like' one another's status updates or photos.
3. A and B once had a friendly collegial relationship IRL. They connected on one social media platform, where each came to realize that the other holds political and social views

that they find objectionable. They disconnect on the platform, and their collegial relationship becomes icier, though it remains respectful.

4 A and B are more or less unaware of each other's existence IRL. However, due to the workings of social media algorithms, they frequently come across each other's posts, which they often find objectionable, on a social media platform that both regularly use.

Scenarios (1) and (2) are those where A and B's relationship remains unchanged despite the possibilities opened up by social media. Scenario (3) represents a case where A and B encounter (what might be perceived as) a lack of care in each other's beliefs that they had not previously been privy to and make a justified decision to disengage. Scenario (4) represents, I believe, the more troubling kind of weak-tie increase that Benkler is concerned with. Scenarios like (4) can, I believe, increase the likelihood of users feeling apathetic toward one another because of the distance (literal and imagined) between them. Without a pre-existing relationship, users may be more likely to view other users as outsiders, even within a single platform (so, aside from cross-platform demographic differences like those described in the previous section), thus exemplifying the pitfalls of parasocial relationships.

Norlock (2017) and Kane (2020) discuss the attitudes underpinning online practices of shaming and accountability, providing an apt example of the ways that online relationships are conducive to apathy. Owing to the *imaginal relationships* described above, social media users imagine themselves to be in particularly arranged moral communities, where they are in community with an imagined audience that expects or encourages shaming and decidedly *not* in community with the shamed (Kane 2020, 595). When someone is shamed online for a transgression of social norms, the shamers come to view themselves as a cohesive social unit; they become members of a moral community on the basis of their shared perception of the shamed party's deserving to be shamed. Even when there are only weak ties (otherwise) present, the discord between shamed and shamer(s) ensures that the groups are encouraged to feel apathetic if not outright hostile toward one another.

Kane's answer to this breed of apathy is a turn toward restorative accountability practices imbued with more "positive social meaning" via an ethics of care[53] (2020, 604). In the context of online relationships, particularly when there is no prior basis for care (that is, when there is no prior relationship that the online element is an extension of as it might be for family members who 'friend' one another online), Kane writes that an ethics of care pays attention to the particular social norms that a space calls for, recognizing such norms as context-sensitive instead of relying on sweeping universal/ist/izing claims. It follows, then, that care ethics might suggest that shaming those we have no prior relationship with is *less* morally acceptable than shaming those we know. That is, the unique contours of online space might make shame less apt of a response. Moreover, Kane notes, a digital ethics of care can recognize that persons' identities are relational, and that persons themselves are *particular* (or concrete, as Seyla Benhabib puts it), others, rather than the abstract moral agents of thought experiments. By subjecting others to online shaming (again, even if shame is warranted), which can easily become disproportionate to the initial offense owing to the involuntary interactions that social media sites foist upon users and can morph from shaming into dogpiling[54], the shamed are denigrated in their status as community members (Kane 2020, 605).

Consider, for instance, a form of online dogpiling that has become especially prevalent over the past decade: that directed toward journalists. Such dogpiling has in recent years been directly called for by prominent political and media figures, and once it has commenced it can be difficult to stimy (Cinnamon 2024, 1; 2). Understanding this dogpiling through Norlock's relational online shaming framework yields this conclusion: the shamers – those contributing to the dogpiling or doxing – find social recognition among one another. That is, they affirm one another. In doing so, they negate (or at least fail to properly recognize, as a member of their

[53] The ethics of care will be discussed in more detail in the subsequent and final chapter.
[54] In a 2024 Berkman Klein Center for Internet and Society white paper, Molly Cinnamon defines dogpiling as "a constant flow of harassing, threatening, and violent messages that utilize and dox personal information online about the journalist, their friends, and their family" (1).

moral community) the shame*d*. Applying Kane's conclusion, "restorative accountability practices on [a social media platform] would involve more meaningful dialog and fewer egregious sanctions," including fewer outright exclusions from moral and social discursive communities (2020, 607).

Before turning to my third and final structural dimension of social media, I want to note that this is not to say that contact with total strangers is necessarily a bad thing on social media platforms. Of course, positive social bonds can arise in online spaces, and being held to account (or 'shamed') can constitute genuine moral education. Therefore, my aim in this section has not been to disparage online relationships in general, particularly given their centrality to young people's social networks; indeed, relationships that begin and solely take place online can be fulfilling, and social media can facilitate genuine friendships. Instead, I have sought to call attention to the ways such relationships can obscure moral boundaries, making it more difficult to practice care and thus rendering individual parties less likely to adopt an empathetic stance toward their interlocutors.

Apathetic Content

Thus far, this chapter has focused on two elements of online spaces that make them especially amenable to an uptick in apathy: the patterns by which individual users come to populate particular platforms, and the kinds of relationships forged by users who self-select onto the same platform(s). We come, now, to the final of the three elements that I discuss: online content. In this section, I will explore this content in two respects: the content itself, that is, the actual material that is shared, and its amplification. Though these might technically be distinct phenomena — the initial decision to post/share content and subsequent decisions to share or interact with it — they cannot be fully extricated from one another, and so I consider them in conjunction. This focus on content is called for because the content shared on social media is "self-disclosed" and "user-generated"; it does not arise *ex nihilo*, and in most cases[55] is shared intentionally (Della Ratta 2020, 102).

To begin, consider an example (mentioned in the introduction to this chapter) that has become fairly commonplace in the literature: political mis/disinformation during the 2016 US presidential election. Misinformation and disinformation are often equivocated, treated as (genuine, true, reliable) information's *others* without sufficient distinction between the two. Sille Obelitz Søe distinguishes the two in terms of their relationship to intentionality; while both represent species of misleading (false) information, misinformation is unintentional whereas disinformation is intentional (2021, 5943). Both misinformation and disinformation are non-natural, Søe writes, because they stand on the same side of the division between states of the world and communications about those states (ibid, 5946). We can infer that there is something about misleading information — or at least about particular bits of misleading information — that make it liable to spread so rapidly. This holds even when that misleading information is overtly antagonistic, and thus when it knowingly or intentionally disincentivizes users' care for groups to which they do not themselves belong.

Because content is designed to be seen, and because it is often intentionally curated (thus the rise in 'content creator' as a legitimate career), it is often designed to evoke a particular affective response. Even more strongly, sometimes *not* capitulating to these affective responses can be a moral failing of sorts. Susan Sontag notes in her analysis of war photos: "[Virginia] Woolf believes that not to be pained by these pictures, not to recoil from them, not to strive to abolish what causes this havoc, this carnage, is a failure of imagination, of empathy" (2002, 83). Photographs, in this case, are taken and shared to stoke empathy, and not responding appropriately would evince a pathological emotional gap. Likewise, this same process can

[55] With obvious exceptions for phenomena like 'revenge porn,' where intimate images are shared without the subject's consent.

encourage *apathy*, either by painting a group in a negative light through photography, or by excluding them from the empathy-provoking photographs altogether, leading to a sense of indifference. This apathy is thus engineered by the 'content' of wartime photography.

Combining these two lines of inquiry — misinformation and crisis-content — allows us to begin crafting a theory of why misinformation that manipulates feelings of em/a/pathy is so common an occurrence in times of crisis. Eriksson Krutrök and Lindgren propose that the "dramatic and triggering character" of emergency-time content allows it to spread as rapidly as it does, even barring factual accuracy (2022, 89). In other words, that information is *highly* emotionally charged (regardless of the precise nature of that emotion) seems to determine that it will spread widely and rapidly. Recall Chapter 1's mention of inspiration porn, the genre of (often digital) content that portrays disabled people as inspirational, worthy of awe, simply for completing mundane tasks "in spite of" their disabilities. Why does inspiration porn go viral? Likely because it consists of one-off cases where a single individual is *so* movingly able to supersede their circumstances, bringing themselves and their loved ones immense joy. Particularly when paired with the assumption that disabled people are incapable of doing anything, these stories get users' attention. The above-mentioned phenomenon of compassion fade again becomes relevant here; even if content is sufficiently affectively-charged so as to stoke compassion (empathy), if users are exposed to enough of it — as they often are on social media platforms where scrolling through hundreds of posts takes relatively little effort — they become less capable of genuine caring responses. I do not want to fully equate compassion fade with compassion fatigue here; the former refers to the decrease of caring responses with scale, while the latter points to a similar trend but with respect to prolonged/frequent exposure to atrocity (see Roberts 2021). Compassion fatigue is often rooted in the inability to sustain 'adequate' care or compassion for the necessary number of victims, while compassion fade is more often presented as a quantitative measure of the utility of interventions. Nevertheless, we can view both phenomena as decreases in individuals' ability to feel the requisite care or empathy that a situation calls for, particularly when coupled with the *distance* (both spatial and [consequently] ethical)[56] that online interactions seem to entail.

A similar pattern is observed with respect to hateful or inflammatory comments made on social media platforms. Part of this has to do with individuals' ability to opt for anonymity in online spaces, but even to outsiders (those who are not actually involved in an interaction) Puryear and Vandello find that face-to-face inflammatory comments stoke more outrage than those made online (2019). This may be because, as they note, "key components of moral cognition" including social cuing, non/paraverbal communication, and the identification of concrete victims "are muted" in such contexts (ibid, 895). This points to another way in which social media content creation and spread are slanted toward apathy: they simply are not viewed as having enough 'to do' with actual, concrete victims, and are subsequently not viewed as victim*izing*. Online content is often shared in a sort of context vacuum. Trying to explain a meme to a friend, colleague, or family member who is 'less online' than they are can have even the most astute digital culturalist sounding like a semi-coherent conspiracy theorist. Meme pages are rife with "absurdist creativity," which almost necessarily entails there being a substantial audience that simply *does not get it* (Phillips and Milner 2017, 93). Online content is polysemous; it can have multiple, even contradictory meanings (ibid, 11). This likely reifies the use-pattern fragmentation identified earlier in this chapter, as well as potentially causing those who are not 'in' on the joke to take satire seriously. If that satire might (incorrectly) be read in a way that promotes apathy or even outright hostility toward a group (likely the group crafting it), then this might discourage empathy or care.

Much of this can be attributed to the work of algorithms, roughly understood as the formulae according to which recommendation systems including search engines and news feeds operate. As Safiya Umoja Noble and Joy Buolamwini point out, algorithms are far from value-

[56] See Slaby (2023).

free mathematical calculations (2018; 2023). They are, instead, human-made, and so retrench the biases of their creators (Noble 2018, 6). Similarly, algorithms are economically responsive; companies can pay to promote their content, so it might appear to be most relevant to (for instance) a user's search when in reality its position at the top of the results list is a matter of its being promoted. So, algorithms are not value neutral. Most algorithms are also proprietary; Twitter does not have insight into precisely how TikTok's algorithm recommends content, for instance, and black-boxing effects mean that it has less insight than expected into how its *own* news feed (or 'timeline') algorithm works. Algorithms often include content that is being shared widely in real-time (see: Twitter's 'trending' topics) without consideration for the veracity of that information or whether it unjustly renders some groups the targets of online vitriol (Webster 2014, 83). Similarly, information on social media has the ability to "go viral" in ways that pre-digital content could not; Webster notes that "the average Facebook user can reach 150,000 friends-of-friends," a magnitude of reach that would take substantially more effort in analog settings (ibid, 42). When content is shared by influencers, those with further reach than such average users, it is not unreasonable to imagine it reaching millions, especially when these high-stakes users are (legitimate or self-proclaimed) experts.

Algorithmic apathy-encouragement is even more apparent in some cases than in others. Noble's *Algorithms of Oppression* begins with an anecdote: a Google search for "black girls" pulling up "HotBlackPussy.com" (an obviously pornographic website) as its first result (2018, 3; 67). This both reflects and reinforces harmful stereotypes about Black femininity and "narratives that reflect historically uneven distributions of power in society" (ibid, 71). Similarly, Noble notes differences in search results for "three black teenagers" and "three white teenagers"; the former populated mugshots, while the latter led to more "wholesome" results (ibid, 81). What these anecdotes reveal is that algorithms are skewed, biased in favor of otherwise-dominant social groups. Noble does not explicitly tie this phenomenon to empathy or apathy, but the connection seems clear: algorithms encourage apathy toward some groups by prioritizing content that presents them as sexually promiscuous and objectifiable, criminal, and otherwise unbefitting of traditional respectability. She explains that stereotypes “create a limited ‘vocabulary of intention,’” limiting the ways that we are even capable of imagining stereotypes’ targets (ibid, 105). If we are limited in how we can see a group, it is more difficult to have empathy toward them as full-fledged members of our moral community.

Conclusion

This chapter has taken a step back from the three previous chapters’ identifying the conditions which necessitate an affective ethics. It has, instead, involved turning toward social media as a site that can make violations of affective ethics — cases where we fail to treat others and their emotions as we ought to — more likely. I have identified three levels of social media infrastructure that incentivize apathy or at least fail to incentivize high-quality engagement. In the following and final chapter, this culminates in a call for attentive care — as an opposite to apathy — in guiding our dealings with others and their emotions.

An earlier footnote gestured toward Jan Slaby’s work. In a recent paper, Slaby argues that structural apathy deserves more attention as the sort of “routine [and socially (re)produced] unconcern” salient in all cases of affective injustice (2023, 65). He writes, further, that:

> Affect is a matter of proximity and distance, and where distance has been created and upheld not just in terms of geographical and geopolitical location, but also in terms of mediatized awareness, public recognition, shared concern, and in the sense of a distance in terms of lifestyle, habit, and day-to-day practice, then affective responses to the suffering or severely curbed life chances of others can be absent or significantly dampened (Slaby, 2023, p. 72).

In other words, affect is a matter of being affected by, and so if there exists a reinforced distance among individuals – as I believe there does on social media platforms – then apathy is presented as an almost value-neutral consequence. To be sure, Slaby is concerned here with apathy that produces or underpins affective injustices, where *injustice* is best understood in the big-picture political sense. His insights are still valuable, though, if we think about apathy as a matter for affective ethicists to wrestle with, especially given that the infrastructural factors identified above can shape how we interact with *any* other users of social media platforms, including those who are situated similarly to ourselves.

These first few chapters have wrestled with violations of affective ethics; with cases where emotions and individuals in their capacities as emotional beings are treated poorly. This chapter in particular looks at the ways that apathy can operate on social media platforms, themselves inherently worthy of consideration because of their increasing prominence in cultural imaginaries. The next and final chapter begins to sketch an answer to the problems outlined throughout this dissertation, making a tentative first step toward an affect-sensitive ethics.

Chapter 5 - Attentive Care Ethics

Introduction

Much of what has been written thus far in this dissertation is critical; of certain emotional behaviors, of theoretical frameworks used to explore the ways that we do harm to others in their emotional capacities, and of the affordances for bad behavior – including bad *emotional* behavior – made by emergent technologies. What is left, then, is to make a positive proposal. In other words: given the affective dilemmas described in the previous chapters, *what are we to do*? What kind of moral outlook – what kind of *affective ethics* – might we adopt to make the answers to the problems identified in the previous chapters clearer? My answer to these questions is the core of this fifth and final chapter. In short, I argue that moral philosophers should concern ourselves with ethical questions of old to meet these (sometimes newfound) challenges: not only *what should we do?* But *how — with what mindset, what internal guiding ethos — should we do it*? Such questions motivate virtually any normative theory, including the one proposed here.

I am unsatisfied with any universalist or one-size-fits-all ethics, one that proposes acting according to some concrete and not-context-sensitive principles. Dispensing with such ethics is among the early care ethicists' chief tasks, as they highlight that different relationships beget different moral responsibilities; I might simply owe something different to my sister or my parents than I do to complete strangers. However, I recognize the utility and allure of such approaches, especially because swinging too far in the other direction might land us in a place from which ethical principles are overly context dependent, thus stripping us of any ability to make generalized claims about the world and how we should be (and what we should do) within it. This is the issue with extreme ethical pluralism. Universal claims are valuable because it is only through making them that ethicists are able to claim *oughts*, to argue for solutions to hypothetical or not-yet-actualized problems. Universal principles exist even when they are not actively being applied. We — ethicists, feminists, denizens of the world — are thus presented with a paradox of sorts: how can we craft ethical principles (necessarily generalized sufficiently to be actually described as *principles*) that maintain some degree of responsiveness to the contours of a particular situation? How can we attend to particulars without rendering our ethics little other than a series of one-off statements?

I propose attentive care — a melding of care ethics as it is usually construed with Iris Murdoch's writings on moral philosophy — as one answer to questions like these. Feminist care ethics, with its focus on relationships and the obligations that arise from them, gives us a substantial part of this answer. We owe different things to different people. What I want to add (or what I want to revive, depending on how strongly you read some of the footnotes and one-off mentions of Murdoch and Weil found in care ethicists' work) is care ethics' grounding in *attention*, the idea that there are better or worse ways of attending to or paying attention to something. Despite the obvious resonances between the two, care ethics texts pay little (if any) attention (no pun intended) to Murdoch and her influences, despite the fact that such texts are enriched by engaging with the idea of (patterns of) attention as moral targets. Attentive care ethics, I argue, highlights moral attention, *good* attention, as a starting point from which we might discern our (other) obligations toward others. In keeping with some of the concerns expressed throughout this dissertation, I want to develop a picture of care as a moral disposition that does not require action, or observable behavioral change. It can remain more strictly affective, in that it can lie in how we see or feel about others, and therefore can accommodate affective wrongs as the intervention point of affective ethics. In developing this picture of care, I underscore care ethics' roots — or at least, the roots I believe the field should recognize — in Iris

Murdoch's just and loving gaze as attending (being *attentive*) to another (1997 [1962], 327[57]). Murdoch's vision of attention (her vision of vision itself, in fact) is particularly useful here because the ability to grow in one's capacity to care, one's readiness to care, without measurable behavioral change helps to address some of the concerns about uptake raised in earlier chapters. In being explicitly *anti*behavioralist, attention or vision (understood by Murdoch as the basis of morality) is fundamentally able to accommodate moral wrongs that are not instantiated in a particular action, thus rendering it able to account for affective wrongs that never rise to the level of uptake (e.g., because they remain unspoken). It can, of course, also give us the tools to describe moral rights – as in, right-doings – which are *a*behaviorist and so cannot be seen through action alone.

I begin this chapter by highlighting some of the foundational texts within the care ethics tradition, beginning with Carol Gilligan's *In a Different Voice* (1982) and concluding with more contemporary writings on care ethics. Care ethics has developed over the last several decades into a moral system in its own right, an alternative to (for instance) consequentialist or deontological ethics, and an entire dissertation would not provide sufficient space to examine even *most* of the texts on it. Nonetheless, I capture some of its essential tenets by looking at some of the field's oft-cited core texts. In the next section, I explore Iris Murdoch's idea of moral attention as a means of making sense of moral change that does not require behavioral change. Then, I combine these two threads, outlining a paradigm of *attentive care* that maintains the core commitments of care ethics while supplementing them with Murdoch's ideas, allowing me to define attentive care — or, in broader terms, attentive care *ethics* — as dispositional rather than behavioral, and so as a helpful moral tool for discerning our obligations toward others' emotions. I conclude by considering what attentive care looks like in practice, drawing on examples from precious chapters. The chapter thus functions in part as a brief history of care

[57] Any citations of Murdoch in this dissertation come from the volume *Existentialists and Mystics* (1997), edited by Peter Conradi. The bracketed year in citations from this volume represents the year that particular essay, interview, or speech was written, conducted, or delivered.

ethics, and in part as a reflection on ways to strengthen the care paradigm through greater engagement with Murdoch and similar thinkers.

Care Ethics

Care in Moral Psychology

Care ethics is sometimes treated as synonymous with feminist ethics because of how the latter critiques gendered divisions of labor, especially *care* labor; that which involves looking after children, partners, the elderly, households, and so on. Care ethics is considerably narrower, though, and it (rather than the broader feminist ethics category, often erroneously conflated with any gender-sensitive ethics) is the target of this chapter. Psychologist Carol Gilligan's empirical work on girls' and women's moral reasoning is often credited as among the first (if not *the first*) foundational texts in care ethics. Gilligan's *In a Different Voice: Psychological Theory and Women's Development* is, at its core, a critique of androcentric moral psychologies. More pointedly, it is a critique of psychologies like Lawrence Kohlberg's six stages of moral development, which are said to culminate in "a reflective understanding of human rights" a moral pinnacle which, Gilligan notes, women and girls seem never to reach in Kohlberg's work (Gilligan 1982, 19). Kohlberg's is an ethic of natural rights and of justice, of doing right by people in accordance with some objective and wide-ranging principles of right and wrong treatment or behavior; Gilligan describes this as the "justice perspective" on moral issues (1982). It "intones the celebration of separation, autonomy, [and] individuation," viewing persons as individuated moral actors rather than emphasizing our connections (and our actions' connections) with one another('s) (ibid, 23).

Kohlberg's six stages are hierarchical; they are not simply six *different* but evaluably equal perspectives on moral matters but instead are arrived at in sequence as a person becomes more wholly morally developed. The further along in Kohlberg's moral sequence a person is — as they progress toward his sixth and final stage — the more morally developed (the more moral, even) they are said to be. Kohlberg's six stages are:

1. The Punishment and Obedience Orientation
2. The Instrumental Relativist Orientation
3. The Interpersonal Concordance or "Good Boy-Nice Girl Orientation"
4. Society Maintaining Orientation
5. The Social Contract Orientation
6. The Universal Ethical Principle Orientation (Kohlberg 1981[58] v1, 16-19).

The moral stages are paired off (1,2; 3,4; 5,6) under more general levels of moral thinking as well. The pinnacle of moral development, the sixth stage, includes "*universal* principles of justice, of the reciprocity and equality of human rights, and of respect for the dignity of human beings as individuals" (ibid, 19; emphasis added). Kohlberg's studies find women often stuck at the third stage at which "morality is conceived in interpersonal terms and goodness is equated with helping and pleasing others" (Gilligan 1982, 18). Kolhberg's six stages are, importantly, increasingly *abstract*, relying on rationale that are increasingly removed from actual happenings. We can see a parallel here to the ways that states must "govern through the creation of universal laws" rather than concerning themselves with particular instances of (e.g.) lawbreaking (Nicholson 1983, 520). Kohlberg's understanding of morality is thus as individuals' internal legal systems, their ways of adjudicating for themselves. If women remain in the domestic sphere – even if they choose to remain there voluntarily – they will, according to Kohlberg's model, remain morally underdeveloped.

Importantly, Kohlberg arrived at his six stages — and consequently at the justice perspective as the highest stage of moral development— through interviews exclusively with boys and men, later writing that he did so "not to add the complicating issue of sex differences to [his] study" (1981 v2, 340). He suggests, even, that "youthful and adult females might be less developed in justice stage sequence than males," and that girls might 'stall out' at earlier stages of moral development, again owing to their likelihood of remaining in the home rather than entering the public sphere (ibid). Gilligan does not take issue with the possibility of boys' and girls' patterns

[58] The stages can be traced back further to Kohlberg's 1958 doctoral dissertation, but the 1981 texts provide a comprehensive overview of his moral psychology.

of moral development differing; her own interviews (that were not limited to males) reveal that any account of moral development that only accounts for boys' moral development is incomplete, not merely for reasons of representation, but because girls and women develop their moral selves *differently* (qualitatively, and not just more or less well) than boys and men. This difference became apparent in Gilligan's studies, where she found that women and girls' ways of "speaking about their own experiences of moral conflict and choice" fundamentally differed from those observed (and subsequently generalized as about *moral development simpliciter*) by Kohlberg (Gilligan 1987/1995, 33). She resists making normative claims about the different kinds of moral reasoning, unlike Kohlberg, rather than claiming that one sort reflects more complete development than another. Gilligan contrasts – rather than ranking – a justice (rights-based) perspective and a care (responsibilities-based) perspective, the latter of which she sees more actualized in girls and women's moral reasoning.

Though the remainder of this section — and of this chapter and dissertation — will focus on care as Gilligan and her successors develop it, it is worth noting that Kohlberg was aware of the gender-based criticisms of his interviews and subsequently-postulated stages. In responding to Gilligan, he writes that "considerations of care and considerations of justice are interwoven in working out resolutions to moral dilemmas," and that "special obligations of care presuppose, but go beyond, our general duties of justice, which are necessary but not sufficient for them" (1981 v2, 343; 229). In other words, our duties to care arise *from* our duty toward justice; care might be a moral principle baked into our conception of justice, and the two should not be thought of in opposition, nor should we think of care as something that we move beyond in pursuing justice-based reasoning. Instead, realizing our duty toward justice is realizing the core rationale for our duty toward care. This meshes with Kohlberg's claims that while Gilligan is concerned with "the differing contents of people's moral judgments, his model is concerned with the more formal structures of moral reasoning" (Nicholson 1983, 523). Theories like Gilligan's, Kohlberg contends, might be ill-equipped to deal with cases where our special (relational) obligations to *multiple* others are more or less identical, but we must deal with those others' competing claims upon us and thus resort to a broader-reaching moral principle (ibid, 229; Joan Tronto reiterates the entwinement of justice and care by noting that the former without the latter is incomplete [1993, 167] as does Annette Baier [1994, 31]).

The Emergence of Care Ethics

Gilligan, inspired by Nancy Chodorow's[59] feminist answer to Freudian psychoanalysis, proposed care ethics[60] as an alternative to Kohlberg's masculinist justice ethics, one rooted in girls and women's moral development after observing that Kohlberg's moral teleology inadvertently precluded girls from being or becoming wholly moral. She posits that, while "male identity [is] forged in relation to the world…female identity [is] awakened in a relationship of intimacy with another person" (1982, 13). Care ethics, then, views persons not as individuals engaged in conflict to ensure that their rights are fulfilled (as a justice ethics might) but as fundamentally "interdependent," as bearing some responsibilities toward one another (Gilligan 1982, 74).

Consider Kohlberg's central example of the so-called 'Heinz dilemma': a man (Heinz) whose sick wife's only hope for a cure lies in drug that is sold at an exorbitant (1,000%) markup, and who (after failed attempts to crowdfund the $2,000 he needs, and failed attempts to negotiate with the druggist) breaks into a local druggist's store to steal the medication for his wife (1981 v1, 12). Gilligan notes that Kohlberg's respondents to the dilemma ground their responses in "a

[59] Linda Nicholson summarizes Chodorow's argument, writing that she "has drawn attention to a culturally universal difference between early female and male socialization: that the first and primary caretaker for girls but not for boys is a member of the same gender as they. One consequence of this difference is that young boys, to develop their own identity as masculine, must negate their early identification with their mothers. As a result, young boys tend to see social relationships as potentially threatening to their sense of self; protection against threats to their sense of autonomy takes on a high value in their lives. Young girls on the contrary incline toward defining themselves in terms of their connections to others" (Nicholson 1983, 517-8). Thus, we can see the male and female ethos as championing autonomy and relationality, respectively.

[60] Daryl Koehn suggests that Gilligan's role in 'founding' care ethics is overstated; she does not deny her importance to the field but suggests that the frustrations that Gilligan identifies in predate her psychological research. She points out, too, that Gilligan never actually defines care (Koehn 1998, 22).

societal consensus around moral values that allows one to know and expect others to recognize what is 'the right thing to do'" regardless of the actual course of action chosen; either stealing is wrong and Heinz acted immorally, or Heinz acted permissibly because the value of human life overrides whatever value is secured by not stealing (1982, 28-9). These responses are often grounded in some sense of Heinz and the shop owner as individuals with equal moral standing; thus, it would be just as wrong (or not wrong, depending on the respondents' views, since Kohlberg focuses more on the reasoning than the conclusion) for the druggist to steal something from Heinz as for Heinz to steal from the druggist. Gilligan notes, however, that her respondents fundamentally view the conflict differently: "[Kohlberg's] Jake [sees] a conflict between life and property that can be resolved by logical deduction, Amy [one of *Gilligan's* participants] a fracture of human relationship that must be mended with its own thread" (ibid, 31). Another way of phrasing this distinction: women "approach moral problems not as intellectual problems to be solved by abstract reasoning but as concrete human problems to be lived and to be solved in living" (Noddings 1984, 96). Gilligan's theory is one that can account for Amy's response as sufficiently moral, on par with Jake's, rather than as underdeveloped, even if to the Kohlbergian it may seem irrational.[61]

Some have critiqued the gendered nature of Gilligan's care ethics, arguing that identifying certain modes of moral thinking as inherently feminine merely reifies the structures that subordinate femininity (Robinson 1999, ch. 2 provides an overview of such criticisms). That is, they argue that establishing any sort of ethical theory along gendered lines is essentializing and prescriptive. We can see avenues toward responding to this criticism in texts like Sara Ruddick's *Maternal Thinking* (1989), wherein the author argues that maternal thinking is a practice, rather than a biological reality. Similarly, Virginia Held writes that we need to focus on care not as an abstract principle, but as instantiated through a series of *caring relations*; practices or

[61] Though Nel Noddings does not quite call the caring relationship irrational, she notes that it is nonrational for both parties (1984, 61).

actions that comport with caring values (2006, 36). As such, it can be performed by anyone, even men, because “a mother is a person who takes on responsibility for children’s lives and for whom providing child care is a significant part of *her or his* working life” (Ruddick 1989, 40; emphasis in original; Angela Garbes’s *Essential Labor: Mothering as Social Change* [2022] provides an even more contemporary look at what it means to build an ethos, a way of living, with mothering as its starting point). Gendered divisions of labor, reinforced by cultures[62] that take them for granted, lead to parenting work being divided between authority (fathers) and care (mothers) (Ruddick 1989, 42-3).

Some argue, too, that Gilligan’s feminist ethic (and, for that matter, Ruddick’s) ignores other identities — racial, class, etc. — that also impact care practices and the assessments of individuals qua carers. Again, the concern here is with essentializing; might it be the case that even if care ethicists do not reinforce a male-female essentialism, their descriptions of caring practices only speak to certain kinds of caregivers’ experiences, thus ignoring (for instance) how women of color’s care practices might differ from white women’s? This is a criticism that Ruddick addresses by writing that she “expect[s] sufficient commonality in the demands made by” differently situated women’s “children to enable us to compare, which also means to contrast, the requirements of our [maternal] work” (1989, 53). That is, if we take mothering (maternality) as the primary moral activity, Ruddick maintains that we can find enough commonality across differently situated women’s experiences of motherhood so as to draw more general conclusions from them.

Nel Noddings’ *Caring* is among the first post-Gilligan texts in care ethics and recognizes moral philosophy’s failures to account for care as failures to include women’s (particularly mothers’) voices. Noddings’ approach to care is more distinctly philosophical (more specifically,

[62] “Women (and men) who feel excluded by the vocabulary of maternity but nonetheless recognize that they think maternally can more easily include themselves in the historically female work of care [when we speak of care generally]. moreover, when it is conceptualized as exploited work, the notion of ‘caring labor’ lends itself to feminist labor politics” (Ruddick 1989, 46). Ruddick nevertheless retains her maternal language because she is speaking of one specific kind of care - that exercised in childrearing.

epistemological) than Gilligan's psychological one. Noddings defines care as a movement away from the self because the one-caring becomes *engrossed* by the cared-for; the importance of this unselfing will become even clearer in the subsequent section. Care, in her view, is fundamentally relational; it does not reinforce individualistic ideals of autonomy. Caring is thus about reciprocity, though not in a Rawlsian sense, because it emerges in caring for and being cared for by others (1984, 14). It is rule-bound without being constricted by reason, without being un- or anti-affective; the rule, it seems, champions recognition or "engrossment," the one-caring's openness to *receiving* the cared-for, and the rule of care has been fulfilled when the one-caring's receptivity of the cared-for is mirrored by the cared-for's response to (recognition[63] *of*) having been received (Noddings 1984, 17; 61; 112). This receptivity can be understood as epistemic. To receive the cared-for is to try to know as they do, to try to occupy the (e.g., cognitive) space they do. Hence the one-caring experiences "motivational displacement," in that they come to desire *as the cared-for* (ibid, 112). So, even if care is enacted — as it often if not always is — asymmetrically or one-sidedly, rather than with one person being immediately able to match another's act of care with an act of their own, it is still reciprocated by the cared-for "achiev[ing] such a state of magnanimous receptivity" (ibid, 76). Noddings' musings about care laid the groundwork for Tronto and Fisher's later contention that care consists of four stages: caring about, taking care of, caregiving, and care receiving (1990; these are expanded upon in Tronto's *Moral Boundaries* [1994]).

If Gilligan's argument for care ethics is psychological (psychosocial, perhaps), and Noddings' is epistemological, Tronto's is best described as *political*. Tronto describes care not merely as something that we instantiate in our relationships with others, "as a moral value" enacted in the private sphere "that should inform politics," but as (more radically) a political practice in and of itself (Tronto 1993, 158; 161). Tronto's definition of care is as "a species activity that includes

[63] This is described in more detail below with respect to Eva Kittay's work.

everything that we do to maintain, continue, and repair our 'world' so that we can live in it as well as possible," and she contends that care practices are necessary for a "just, pluralistic, democratic society" (ibid, 103; 162). Such practices obviate (inter)dependence, the connection between 'interests' (often viewed as individualistic) and 'needs' (viewed as of public import because of their priority) and sustained moral engagement as fundamental human values necessary for healthy democracies. Moreover, Tronto writes, a care-ful polity melds the public and the private, making even some areas that have previously been classed as the domain of autonomous individuals into objects of political concern. Others including Fiona Robinson also propose a political ethic of care. Robinson posits that, insofar as care is linked with need (which is itself linked with vulnerability), care ethicists cannot in good faith ignore the ways that certain populations are rendered need-y, through the theoretical lens of human security. Care, then, is not a politically or normatively neutral category, and seeming limitations to care (e.g., the idea that it can only be enacted amongst proximal parties) are not reasons to abandon it, but to optimize it (Robinson 1999, 48-9).

Lastly, many feminist care ethicists (and, indeed, critics of care ethics) have focused on the elements of a *complete* ethic of care, an ethic grounded in *completed* caring actions, writing that we cannot champion caring practices without focusing on their completion. Eva Kittay writes: "'care' is CARE[64] only when it is completed through its reception" (2019, 186; more on the application of care in disability contexts comes below). That is, an act of care is only properly the kind of moral care feminist ethicists target when the one cared-for accepts it, thus accepting the broader relationship that acts of care indicate. Kittay considers, too, the myriad ways that disabled people (particularly those with profound cognitive disabilities) can receive care, as many who are cared-for might not be able to outwardly thank the ones-caring. Care fundamentally requires a response on the part of the cared-for (ibid, 194) This makes care

---

[64] The capitalized CARE denotes "open responsiveness to another," rather than mere regard, which care is sometimes colloquially used for (Kittay 2019, 7.2).

reciprocal in two senses: as mentioned above, it is reciprocal in that the cared-for and carer roles can switch, but it is also reciprocal in the sense that an act of care might be incomplete until it is received by the cared-for.

This is far from a systematic overview of care ethics and its applications. Such an overview would span multiple dissertations of this length. Instead, what I have shown in this section is that an ethic of care is a relational ethics, recognizing that different circumstances — where circumstances can include individuals' standings with respect to one another — beget different moral obligations.

Murdoch's Vision

The 'attentive' in the picture of attentive care I am proposing is rooted in Iris Murdoch's vision of just and loving attention as a moral disposition. Murdoch develops just and loving attention in response to (the twentieth century's) predominant trends in moral philosophy, according to which "the moral life of the individual is a series of overt choices which take place in a series of specifiable situations" (1997 [1956], 77). This represents a distinctly *behaviorist* ethic, according to which persons' moral character is assessed on the basis of what they do. Morality is thus about "acts and choices," which are thought to be observable; moral statements about the goodness of an action are commands that particular things be done (1997 [1956] 79; [1957], 63). This view — which Murdoch dubs the *current view* of her age — also maintains that moral concepts consist of pairings of facts and recommendations, and that moral judgments are universalizable. Consider, for instance, utilitarian and deontological ethics, both of which propose objective, universalizable ethical truths rather than making these truths context dependent. In disputing strict behaviorist[65] ethics, Murdoch writes that we need to blur the line between inner life (thought, vision) and behavior — a line that has historically been underscored by moral philosophers — shifting the metaphorical language that we use to talk about morality

[65] Murdoch also critiques the predominant ethic as being existentialist "in its elimination of the substantial self and its emphasis on the solitary omnipotent will" [to action] and utilitarian in its preoccupation with public — observable — acts" (1997 [1962], 305).

toward one grounded in seeing or attending to (1997 [1962], 302). Attention, for Murdoch, is not merely a perceptual ("impartial or trans-historical" [Bowden 1998, 59]) capacity, but is something that can be normatively evaluated. We can attend more or less *well* to a situation.

This is best illustrated through Murdoch's chief example of a woman (M) and her daughter-in-law (D), briefly mentioned above. M "feels hostility to" D, finding her "pert and familiar, insufficiently ceremonious, brusque, sometimes positively rude, always tiresomely juvenile" (1997 [1962], 312). M, as a woman who generally cares about behaving properly, "behaves beautifully to the girl" despite her misgivings about her character (ibid). Because M "is an intelligent and well-intentioned person, capable of self-criticism, capable of giving careful and just *attention* to an object which confronts her," she critically interrogates her own feelings toward D, and "gradually her vision of D alters" (ibid 313, emphasis in original). What is crucial here is this: nothing about M's behavior toward D has changed, but her internal assessment of her has. Murdoch resists the impulse to classify M's changing ways of seeing D as "mere nothings," but instead as changes in M's attention: "M *looks* at D, she attends to her, she focuses her attention" (ibid 314; 317, emphasis in original). M comes to see D, trains herself to see D, "justly or lovingly," where "love is knowledge of the individual" (ibid 317; 321).

The change in M is a moral one because it represents her coming to direct her attention on D in such a way that represents her directing her attention toward the good. Murdoch draws inspiration here from Simone Weil, whose religious account explains the primacy of the faculty of attention because of its conduciveness to prayer. Attention, writes Weil, represents desire, and "consist of suspending our thought, leaving it detached, empty, and ready to be penetrated by the object" upon which we are turning it (1952, 59; 62), Weil maintains that this attention, turned upon the proper object — God — is the essential character of prayer. Notably, too, attention is effortful; this seems to be, for Weil and Murdoch alike, what makes it a moral achievement rather than simply a state of being. Attention, in the moral sense, is not mere transfixation on an object in order to decode it (in Murdoch's example, it would not merely consist of M focusing all her attention on D) for the purposes of interpretation but is a more general principle of seeing *well* while maintaining an essential emptiness[66] of thought (1952, 62). Seeing the Good as well as the good pray-er sees God is the crux of just and loving attention for Murdoch.

Lastly, the attention — toward God, toward the Good, toward whatever we want to justly and lovingly direct it — must be a kind of un-selfing.[67] Otherwise, we risk our purported just and loving attention being little other than egoism. Attention must be focused, and it must be directed toward some object that is external to the self. Otherwise, we see the other not as they actually are, but as we wish (maybe for some self-serving reasons) to see them. Likewise, the failure to unself can represent a sort of social convention: "because we are ourselves sunk in a social whole which we allow uncritically to determine our reactions," (Murdoch 1997 [1959], 216). This is why, for Murdoch, art — especially literature — provides an especially apt medium through which to hone our attention. Literature facilitates our unselfing because it presents us with a reality within which we can become engrossed, but which the egoism of our everyday lives might otherwise obscure (Bowden 1998, 64). When we are sufficiently engrossed, our 'I' (or perhaps our 'we) disappears.

A straightforward critique of Murdoch might say that she asks too much of us, especially given our dealings with morally reprehensible people. In other words, by instating a morality based in just and loving attention, is she suggesting that we need to *love* even those who harm us or others (or who seek to)? In short, no. To gaze upon them lovingly is to see them accurately, though, and does not make any prescriptive statements about how we are to evaluate what we see or how we should treat the object of our gaze afterward. Gazing lovingly, paying just

66 Another helpful example from Murdoch here is that of her turning her attention toward a kestrel.

67 Though Murdoch says this explicitly, there is debate over the extent to which it should be read backward into Weil's work. Peta Bowden says that the commitment described by Weil is actually self-reflexive, in that it rests on our granting "the contingency and inappropriateness of 'lookings' that are formed through our personal attachments unveils our participation in the order of impersonality and truth" (1998, 62). There is also a dialectical sense in which recognizing the fallibility of our own perceptual capacities is inseparable from recognizing that others may turn their (similarly flawed) capacities upon us.

attention, means understanding persons in their contexts, and seeing them "with goodwill" (Carrasco and Alegría 2023, 61). Understood in this way, I see Murdoch making a similar argument to theorists who suggest that we ought not view even the most morally reprehensible persons as monsters, because by doing so we fail to recognize them as people and, by extension, present their evilness as something other-than to humanity. There is a similar critique of care ethics holding that it is too demanding an ethic and asks us to respond in a particular way to cases where "what we hear from the other arouses feelings of alarm, disgust, or doubt" (Noddings 2010, 9). This meets a similar answer: that caring is about maintaining a particular sort of disposition toward the other, a commitment to seeing them and their needs and wants, and not about maintaining any specific normative or evaluative judgments of them.

I have chosen to ground the version of attention I focus on in Murdoch's work rather than that of two other contenders — Simone Weil (mentioned above) and Martha Nussbaum — but I cannot understate the importance of their ideas about ethical attention. As mentioned, Murdoch explicitly locates the roots of moral attention, of attention as a moral capacity, in Weil's work (1997 [1961], 293). I find Murdoch's work more applicable than Weil's outside of religious contexts, though it is important to recognize its roots. Nussbaum's account of moral attention, too, is helpful for making sense of Murdoch's. Nussbaum begins a 1985 paper with a bold assertion: that "our highest and hardest moral task is to make ourselves people 'on whom nothing is lost'," meaning that *seeing* (being attentive) and resisting obtuseness is our primary means of acting well (1985, 516). This kind of seeing, Nussbaum contends, involves both generalities and particulars: we see against the backdrop of a set of background commitments (generalities), but we use those generalities to shape our understanding of a particular situation. Likewise, the particularities can be used to turn backwards and reshape our general commitments (Bowden 1998, 68; Nussbaum 1985, 524; 526). What makes Nussbaum's theory critically different from Murdoch's, though, is that the former's is grounded more in an egocentrism — not understood in the negative normatively-laden way in which the word is often used — because of this reflexivity. A large part of what is important about moral attention for Nussbaum is its ability to make us think about ourselves and our values, and thus to shape our sense of duty (cf. 1985, 525), whereas for Murdoch the self ideally falls away in our concerted focus on the object of attention. I find this unselfing aspect of Murdoch's especially vital, given how often moral decision-making asks that we set ourselves and our interests aside. In the words of an oft-memed Instagram story by actress Jemima Kirke: "I think you guys might be thinking about yourselves too much."

Attentive Care: An Ameliorative Disposition

We come now to the point where the previous two sections' contributions to the literature can be synthesized. Care ethics is no doubt valuable insofar as it is a contextual ethics, rather than one which prescribes a single principle according to which we should make decisions. Moreover, the reciprocity of care (highlighted in Kittay's work on the *completion of care*) makes it compelling for interactional ethics. There is a colloquial resonance between Murdoch's work and care ethics because it seems impossible to care about something without paying attention to it. The sentiment is so trite, in fact, that it's mentioned in *Lady Bird* (Gerwig, 2017), when a nun asks the protagonist: "Don't you think maybe they're the same thing — love and attention?". Melding care ethics with moral attention in a more intricate way, I contend, allows us to craft a non-action-constrained ethics that can thus exist in an unabstracted way without requiring uptake of some action or speech (because it subverts the action-requirement, and actions are what gets taken up). This section first explores the ways that care ethics already makes use of language about moral vision before turning toward a richer theory of attentive care that shows the possibilities enabled by combining these threads of thought.

Some — though quite few — writers have called attention to the resonances between Murdoch's thought and that of more contemporary care ethicists. Claude Gendron notes that this work has primarily come in fields other than moral philosophy, especially psychology (2016). Nonetheless, care ethicists' explicit concern with Murdoch's work has been quite limited, which I find surprising. This point is reinforced by Noddings' claim that "[m]oral philosophers," with the noteworthy exception of Weil, "have not said much about attention" (2002, 14). I would contend that Murdoch is the clearer exception than Weil, and that her work is the theoretical bedrock upon which most of contemporary care ethics sits. Noddings acknowledges (though briefly) Murdoch's just and loving gaze as one strengthened by practices of caring in the afterword to later editions of *Caring* (2013), and in a 2010 article about caring in education in which she explains her shift from talking about engrossment to attention (2010, 8). The most comprehensive account of Murdoch's preempting care ethics comes from Carrasco and Alegría (2023). They note two distinct features of Murdoch's writings that make it relevant to care ethics: her focus on moral attention (as a feature that can be honed over time), and her emphasizing individuals' particularities (Carrasco and Alegría 2023, 61). The authors seem to evade any deeper consideration of Murdoch's grounding care ethics, addressing the gaze as the starting point for ethics of care. Delving deeper into this resonance is precisely my task below. I accomplish it by combining the threads — care and attention — in two distinct ways: first by looking at ways of seeing that are informed by care ethics' core tenets, and second by thinking about how care might be reshaped given its grounding in Murdoch's thought.

Below, I consider what moral vision or attention might look like within existing care-ethical contexts. Then, I more fully explain *attentive care ethics*: a version of care ethics informed by the flavor of moral attention just mentioned. I consider, as I explore in these and in the final section below, attentive care to be an ameliorative affective disposition. That is, it is a non-action-based ethic of seeing.

Attentive Care: Seeing in Care

Stepping outside of oneself as care ethics asks us to is necessarily about vision. Thus, even though few care ethicists make direct recourse to Murdoch (or Weil), we can see them discussing moral attention in some capacity. This is particularly true insofar as care ethics requires tailoring action to specific interlocutors (so that acts of care can be adequately completed), rather than making sweeping claims about (e.g.) maximizing utility irrespective of particularities. Nel Noddings explicitly links vision to care ethics using an example of a mother (as the one-caring) teaching her child how to regulate his emotions: “The one-caring receives the child and views his world through both sets of eyes…The one-caring assumes a dual perspective and can see things both from her own pole and that of the cared-for” (1984, 63). A crucial part of caring in this case (and, as we’ll see, in others) is double-vision in the sense of the one-caring’s seeing and expressing from their own point of view and from that of the cared-for (ibid, 60). If either of these perspective-takings fails, the cared-for either becomes completely subsumed (and thus erased) by the one-caring or the one-caring loses their *own* agency in their subserviency to the cared-for. In this case, seeing with care entails seeing the present (in Noddings’ example, a child’s unhappiness after a bad day at school), and seeing a way toward a better future.

Above, I have mentioned the notion of care being completed upon its reception. This, too, requires a kind of vision because it requires that the ones-caring “see the cared-for as he is and *as he might be* “(ibid 67, emphasis added). Though the ethics of care is grounded in practices of care, care itself involves envisioning the future in relation to the present, and envisioning ourselves playing some role in securing that future. This is why care is sometimes ‘mere’ maintenance rather than (for instance) acts of restoring an ill or aggrieved loved one to their prior emotional stasis (see Fisher and Tronto 1990, 40). In other words, some acts of care are accomplished by seeing the world as it is and endeavoring to keep it (looking) that way.

What makes attentive care especially effective, in my view, is its ability to stand well as a moral disposition, rather than simply a principle of action. If we think of utilitarian or deontological ethics, they both seem to require action. This does not mean that those who live according to these ethics need to be acting constantly, but that the ethics themselves do not become relevant until the moment of action. Murdoch’s critique of behaviorism makes clear that her predecessors, in focusing too extensively on action, leave us with “far too shallow and flimsy an idea of human personality” by abdicating with moral concepts (1997 [1961], 287; 290).

Attention, on the other hand, operates fundamentally on the internal; it might get expressed later through action, but its expression — through actions that show what a person has attended to — is secondary. Joan Tronto actually critiques the notion of care as a disposition, writing that to view it in dispositional terms "allows us to think of care as the possession and province of an individual" (1994, 118). In her view, this reinforces gendered stereotypes about care as women's domain, and about women being properly contained to the realm of the private. Her concerns are certainly worth raising, as they echo other criticisms of care ethics as gender-essentialist. I think these concerns are still amenable with a (proper) practice of attentive care, though; note that Tronto's concern is with those who view "caring [as] *not so much about the activities of care*, but about the emotional investment that has been made in order to care" (ibid). I do not view these two possibilities as incompatible. Though I have mostly been talking about attentive care as something that does not require action, this does not mean that it *cannot* be enacted. It can and often is. The benefit of attentive care is that it allows us to account for morality — for moral *change* — before or irrespective of action.

Moreover, Tronto's 1994 expansion of she and Fisher's four elements of care makes clear the role of vision in well-done care. In this later work, Tronto writes that the four elements of care (caring about, taking care of, caregiving, and care receiving) generate four *ethical elements* of care. The first element of care is caring about, which begets the ethical element of *attentiveness*, namely attentiveness to others' needs (Tronto 1993, 127). Here, too, I think she speaks in terms that are a bit too reductive: she notes that Weil views "the absence of attentiveness [as] a moral failing," but fails to acknowledge Murdoch, who discusses attentiveness in a way that it more directly connected to care ethics (ibid). Accordingly, the picture of attentiveness she leaves us with is not as nuanced as it could be because it does not seem to account for better and worse ways of seeing, of paying attention. This is a benefit of using attention in the Murdochian sense I have been wielding it in, as Murdoch resists the urge to flatten attention into something that is either being practiced or not.

Attention, too, insofar as it represents an openness, helps us move beyond dyadic structures of care according to which the cared-for and one-caring are clearly delineated in particular actions. This openness is inherent to all the offered understandings of attention — Weil's, Murdoch's, and Tronto's — alike. For example, consider Murdoch's classification of attention as "a just and loving gaze directed upon an individual reality" (1997 [1962], 327). Her use of *an individual reality* rather than *an individual's reality* is telling here, because within an individual reality we can attend to multiple different individuals' realities.

Because it is dispositional rather than fundamentally reactive, attentive care can help to alleviate some of my earlier concerns about non-uptake-based injustices being overlooked. This is among its primary benefits. In other words, adding attention to care allows us to theorize care without some attendant theory of action. Some early care ethicists might seem to acknowledge the problem of care without action — Noddings, for instance, asks "Is direct, externally observable action necessary to caring?" (1984, 10; elaborated on p. 23). It is worth noting, though, that she seems to approach non-action caring as a matter of *choosing not to act*, of choosing to withhold (presumably some particular) action rather than (as I do) as a matter of action never being a considered or live option. Murdoch's study of M and D provides a clear example where a person's attention is being redirected — and in a way aligned with love or care — but there is no requisite behavioral shift. I expand upon this in the following section, reconsidering examples from previous chapters in light of this enhanced ethic of care.

Attentive Care in Practice

Attentive care, as almost any moral framework, makes most sense in context. In this final section, I take up the phenomena in the previous chapters of this dissertation — morally condemnatory feelings that are unaccounted for by the affective injustice paradigm, the specific breed of affect alienation that arises from the application of unjust social norms pertaining to the treatment of emotions, and social media structured so as to disincentivize care — through the lens of attentive care. I focus especially on the *attentive* of attentive care, given the breadth of established literature on applications of care ethics more generally. Below, I argue that a disposition of attentive care: (1) Allows us, as mentioned briefly above, to prevent the oppression that occurs via unjust emotions, even when they remain unexpressed; (2) Calls for sufficient reciprocity to clarify the limits of our obligations toward others' emotional expressions by creating a general obligation to see accurately without enforcing an obligation to engage; and (3) Can be applied at the infrastructural level as a disposition that lessens the culture of apathy prevalent in online spaces.

Bad Feelings: Attentive Care Ameliorating Ableist Affects

The first chapter of this dissertation addresses the emotional dimensions of injustice by looking at attitudes manifest in the ableist imaginary. The three identified emotions are pity, fear, and awe, all of which can be overtly weaponized against disabled people, or can function as harmful against disabled people even when they remain unspoken owing to their role in maintaining an imaginary in which disability begets pity, fear, and patronizing awe, foreclosing the potential of positive counterparts to these affects. The following chapter made apparent that the extant affective injustice literature is unable to account for the wrongfulness, the injustice, of these emotions *when they remain unspoken*. Authors writing in this tradition have focused on unjust uptake, whether that is a subject's emotions being taken up improperly, or a subject's being forced to metabolize excess or prejudicial uptakes from without. In both of these cases, the 'subject' is the one who is targeted by the affective injustice.

Perhaps an unspoken weakness of the affective injustice paradigm lies in authors' resistance to adopting a specific ethos. That is, staking their claim as (e.g.) virtue ethicists or care ethicists seems beyond the scope of most writing on affective injustice. More explicitly adopting an ethic of attentive care might, however, make theories of affective injustice more amenable to non-actionable injustices; injustices that need not be expressed in order to 'count' as unjust. What impact would adopting an ethic of attentive care have in the situations described in the first chapter? In short: because attentive care directs not only how we act but how we see the world, it would curtail the negative attitudes toward disability that often go unchecked.

We can break attention into two dimensions: seeing others as they really are and seeing them as they wish to be seen. To 'see' poorly, without proper attentiveness, means to see not only wrongfully but wrongly; it is to see inaccurately. Both axes of proper attention entail unselfing, moving beyond a form of perception that is solely self-motivated or -directed. Seeing poorly (with respect to the earlier chapter's examples) would mean seeing either in self-interested or self-promoting ways or seeing inaccurately altogether. Pity, fear, and awe as described in the first chapter thus represent poor (moral) vision. Such affects, too, contradict our responsibility to care for others; to *care about* how they see themselves and wish to be seen and the narratives they wish to be party to. Quite importantly, these feelings do not need to be taken up to exist as wrongful, as they do not need to be spoken at all. Therefore, rectifying or correcting them might have no behavioral or outward evidence, as with the shift (in Murdoch's example) in how M sees D.

Recall the claim that pity suggests deflated social status; we assume that those who we pity are altogether worse-off than we are. Especially when disabled people are the objects of pity, this is a fraught assessment, given that our imaginings of what disabled lives (and the *quality* of these lives) are like can be quite misguided. Disabled people, by virtue of their "alternative ways of being in the world," might judge their own quality of life according to principles and measures that might seem totally incoherent to the nondisabled, and particularly to the nondisabled *and ableist* (Kafer 2013, 83). Or, even when disabled people use familiar metrics to assess their own quality of life, there can be differences in their self-assessed quality of life and that which others estimate, with the latter often falling short of the former. Therefore, assumptions about quality of life — and assumptions that living with a disability leads to poor quality of life — are often grounded in poorly attending to disabled people's actual lives (and the quality thereof).

A similar logic holds for fear. As a reminder, the fear I discuss in a previous chapter is not necessarily fear of becoming disabled but is fear of disabled people themselves. The visual aspect here is perhaps clearer than it is with pity; being afraid of disabled people involves seeing them as something other than what they are, as monstrous in some way. Alternatively, it can involve seeing them as a reflection of one's (the seer's) own frailty or mortality. Seeing in these ways fails to meet the conditions of attentive care, because it is inaccurate, at odds with how disabled people wish to be perceived, and/or egocentric. There is no reciprocity, identified above as a key element of all ethics of care, here; one party is fearful, and the other is made into the object of their fears.

Lastly, we can read ableist awe - which often looks like the patronizing imposition of narratives about overcoming — through the same lens. To be sure, disabled people may at times like being the object of awe; it can feel good, particularly if done to recognize an accomplishment they often feel goes unrecognized, or if the awe is not at their outsmarting their disability but their pursuits despite rampant structural and interpersonal ableism. But there are cases, as expressed in chapter 1, where awe is oppressive. These cases, like those mentioned above, involve a disabled person becoming party to a narrative that they did not consent to, and that they (often) do not find an accurate description of their world. This includes so-called inspiration porn. Moreover, they turn the disabled subject into a tool through which the awed person can project some assumptions about *other* disabled people (e.g., that they are not working hard enough and so do not rise to the same level as the awe-worthy one).

If these injustices can be understood as failures to see (*well* or in some cases *at all*), then corrective moral vision — the attentive vision concomitant with attentive care — can serve as an ameliorative. Attentive care is reciprocal, rather than rendering one party useful for another but not vice versa. In being reciprocal, it involves the one-caring[68] receiving the cared-for as they

[68] I do not want to suggest that disabled people cannot be caregivers. I use *one-caring* and *cared-for* here to refer to the fact that the ableist gaze is the one that is being changed. This gaze, ideally, becomes that of one-caring, one-actually-attentively-caring.

actually are, rather than how they wish them to be. Adopting a disposition toward attentive care would certainly lessen the prevalence of ableist pity, fear, and awe, and of the ableist imaginary more generally.

Attentive Care and "Therapy-Speak"

The subject of this dissertation's third chapter is the opaque nature of our obligations toward others' emotions (often expressed through emotion-responsive norms), and how this opacity can make affect aliens of those who fail to be sufficiently responsive to social norms surrounding emotions. There are different levels of emotional marginalization and psychiatrization at play, all of which are at least in part amendable by adopting a disposition of attentive care. In this subsection, I address both the kind of psychiatrization that leads to anger being re- and mis-read as Madness (that is, as psychiatric ill-health), and to the psychiatrization inherent in the uncritical application of so-called 'therapy-speak' as evidence of socially compelled deference to others' emotions.

Recall from the earlier chapter that members of marginalized groups — whatever axis of identity they are marginalized according to — are particularly liable to having their emotions misread. This can be a matter of scaling, like reading mild frustration as outright rage, or can involve a more complete affective transformation, where an emotion is taken up not as a stronger version of itself, but as something else altogether. The primary example of this in the third chapter was when anger gets taken up as Madness, as evidence of psychiatric illness or deficiency. How might this represent a failure of attentive care? One answer to this lies in the accuracy of just perception: to see someone as having a disorder by virtue of their emotional expressions might not be truthful. Attentive care, then, calls for seeing persons *including their emotions* as they really are, to the best of our ability to have a sense of how they really are. This answer might run into issues, though, if we think about those who actually have diagnosed psychiatric disorders — those who are, in fact, Mad — and whose apt, righteous emotions are dismissed, regardless of whether this dismissal has to do with the diagnosis (which may hinge on how much information the person making the accusation of Madness knows about their interlocutor). Even in these cases, I want to claim that the standards of attentive care have not been met. Justifying this claim requires a bit more of an in-depth look at writings on care ethics and psychiatric disability.

It is beyond the scope of this project to gloss the entire literature justifying the inclusion of the mentally ill in our ethics;[69] I take for granted that they are persons and should be respected as such. So, the next step is to ask what it means to enact an ethic of (attentive) care with respect to the mentally ill, the psychiatrically disabled. I would argue that this involves, in part, seeing them as they are (as mentioned above) with the added corrective of recognizing that there are prejudicial cultural assumptions attached to *how they are*. That is, even attending to those who *actually* have psychiatric illnesses (rather than only those who are accused of having them on the basis of other identity-prejudices) means making an active effort toward seeing well. In a case where we are faced with a psychiatrically disabled person's emotions, and we believe (for non-prejudicial reasons) these emotions to be disproportionate or otherwise inapt, sometimes attentive care will require that we accept these extreme emotional states as meaningful, even if not reasonable, in the same way that even "sane" (or non-mentally ill) people can act unreasonable (Radden 1985, ch. 5). In some cases, therefore, "defects of reason" — unreasonableness — can be taken to justify excusing mentally ill persons for their seemingly inapt emotions (ibid, 100). In short, an attitude of attentive care toward those displaying emotions that might get them classified as insane involves (1) a reassessment of the person's emotional state, with a particular eye toward the ways that social prejudice might be clouding our judgment and (2) if the former leave us still questioning the aptness of an emotional display, recognition that some people are vulnerable to unreason. I recognize that this is an unsatisfying answer, in part because of the difficulty we face in non-prejudicially evaluating an emotion for

---

[69] Jennifer Radden's work is particularly illustrative here, as are Eva Kittay's writings on the moral status of the cognitively disabled.

its aptness. Nonetheless, at a minimum we can say that we owe some level of attentive care to mentally ill persons who display inapt emotions, in part because what it means to attend to them well (genuinely, justly, lovingly) is to recognize their vulnerability to such emotional displays, recognizing such vulnerability as a reason for further inclusion in the moral community rather than exile from it. This can be understood as an attentively-caring obligation to *see* without an obligation to *entertain*.

A portion of chapter 3, too, focuses on the phenomenon of (restorative) therapy-speak: the platitudinous importing of psychological concepts into everyday (non-clinical) conversation. Like psychiatrization, problematic applications of therapy-speak can be understood as failures of attentive care. Problematic applications of therapy-speak include cases where it is used to convey (however indirectly) an obligation of deference toward others' emotions. Although such deference is often well-intended, and can be morally praiseworthy under some circumstances, it represents a failure to see individuals as such.

Recall that the care ethicists' and Murdoch's writings involve attending not just in general but attending to the specific situation and person(s) in front of us. In Murdoch's paradigm case, M was not merely practicing (what is often referred to as) mindfulness; she was reevaluating how she attended to her daughter-in-law in particular. Thus, the salient part of attentive care for addressing issues like those raised in chapter 3 is its focus on (what Seyla Benhabib calls) *concrete* others. Benhabib contends that dominant moral theories (those that Murdoch critiques) view the moral self as "a *disembedded* and *disembodied* being" (1987, 81, emphasis in original). Though this may at first glance seem to mesh well with Murdoch's idea of unselfing, it is actually at odds with it; unselfing in Murdoch's sense requires that the (un)selves remain embedded because they must be able to perceive, to direct their just and loving gaze upon others. Benhabib and Murdoch's critiques of dominant moralities are therefore compatible. In response to her own concerns about generalized (disembedded and disembodied) others as the sole subjects of moral theory, Benhabib proposes the *concrete other*. She recognizes that the generalized other has a role to play in theories of care, as it is in relation to such others that we make sense of "norms of *formal equality* and *reciprocity,*" but adds that the concrete other involves "view[ing] each and every rational being as an individual with a concrete history," and that our relationships with concrete others are "governed by the norms of *equity* and *complementary reciprocity*" (ibid, emphasis in original). These are the norms that primarily govern our interpersonal interactions.

Applying therapy-speak-isms uncritically represents treating individuals as unindividuated, as generalized rather than concrete others. Accordingly, then, it can be read as a failure to attend to individuals as such, and insofar as attentive care requires attending to individuals applying therapy-speak in these ways is a failure of attentive care. Moreover, insofar as the norm dictating the aptness of therapy-speak is a social norm — that is, it is a norm about how we ought to respond to others' displays of emotions — holding tightly to such a norm might convey a disregard for neurodivergent people and others for whom these norms are opaquer.

Practicing Attentive Care on Social Media

The fourth chapter of this dissertation focused on social media as a site of affective injustices, exploring three distinct (infra)structural layers on which platforms incentivize apathy, thereby disincentivizing care. In it, I focused on individuals' self-selection into social media repertoires based on platform use, the kinds of relationships that social media users tend to form, and the exact nature of the content shared given that affectively strong content tends to spread more readily. In what follows, I argue that, though these features are not themselves 'bad' (to avoid moral oversimplification and to recognize that they might contribute to platforms' and tech companies' goals) they make it more difficult to practice attentive care. I do so by, as in the previous chapter, considering the three structural levels separately.

The first of the three levels I considered in the previous chapter has to do with individuals' choices about which platforms to use (and thus to add to their personal social media repertoires). Such choices tend to fall into patterns with some baseline demographic similarities — however general — among users of a single platform. These demographic trends carry from platform-to-platform, and thus at the level of media *repertoires* rather than remaining confined to singular platforms. The connection with moral vision here is clear but exists in a way that has perhaps been understated in the earlier parts of this chapter. Recall that, for Murdoch, the just and loving gaze is an act of moral attention turned upon a particular object. What this entails, though it may seem so self-evident as not to be worth saying, is that these objects must be within our vision to begin with. That is, you cannot work to see better that which you literally cannot see to begin with. Therefore, moral attention might require exposure or coexistence within the same publics, which might require less stringently bounded, polarized social media platforms from within which users are not exposed to 'others' to begin with.

This answer is not wholly satisfactory, though, because it might appear to neglect our moral obligations toward those we cannot see. This is a problem that care ethicists have historically been aware of, as a common critique of certain limited ethics of care rests on their (purported) inability to make sense of moral relations with those with whom we are not in close proximity. I find the political ethicists of care — Tronto and Robinson, for instance — best-equipped to respond to this concern, though Robinson roots it in earlier feminists' worries. For instance, just as distance may render us totally unable to care, it might invite colonializing forms of care (Robinson 1999, 44). Relation across distance is a central benefit of social media platforms, as most uses of them do not, strictly speaking, limit users to interacting with those who are geographically near. So, how can attentive care help us strike a balance between disregarding distant others no matter how severe their plight, and looking upon them patronizingly?

The answer here lies in two key features of attentive care: that it requires attending to reality as it actually (*really*) is, and that it can be dispositional rather than requiring discrete action. Maintaining an attitude of attentive care toward geographically distant others, or toward those rendered distant by a lack of overlap in social media repertoires, does not require that we take caring action toward them, but only that we maintain a commitment to seeing them as they really are *if we actually ever see them.* Though attentive care can facilitate a wide range of ethical actions, it strictly requires relatively minimal commitments, because it dictates a way of (actually or hypothetically) seeing. Therefore, even some of the alienation caused by self-selection into social media repertoires can be minimized so long as users maintain a disposition toward attentive care when outside of their chosen platforms.

The second level of analysis in the previous chapter pertained to the kinds of relationships that social media platforms tend to cultivate. There, I focused on social media platforms' propagating ties (relations) that are greater in number but not in strength or quality. In fact, weaker ties can encourage apathy (and disincentivize or even outright discourage care) because, as mentioned above, care is harder to sustain across distance, including the sort of ideological distance that we envision between ourselves and those with whom we are acquainted yet not friendly, or even more so those whose 'takes,' whose content, we are exposed to only in passing. It is easier in situations like this — where you might only see a post or two of a person's — to see them poorly, because the timeline structure of social media platforms does not encourage (or often even allow) us to see persons in their entirety. In cases like this, as with those where repertoire-selection creates similar antitheses to care, attentive care as a disposition is important. It helps us maintain awareness that even those who we are exposed to only in passing have rich, complex inner lives.

At this point an idea mentioned above bears repeating: attentive care does not entail any particular normative evaluation. That is, we can care-fully attend to someone and find them (to use perhaps extreme language) morally despicable, provided that this is in alignment with what our attention reveals to us. So, just because weak ties make it more difficult to view others with care, coming to accomplish such a view does not mean that we have to end up with any particular view of them, other than viewing them as they really are.

Lastly, the third section of the previous chapter examined the content shared on social media and the algorithms that amplify some posts while rendering others less visible. In that section, I noted that posts that convey extreme emotions (recall Sontag's essay on responses to wartime imagery) tend to garner more attention. This leaves those that do not make explicit emotional appeals — or those that appeal to emotions that are less compatible with public opinion — receiving less visibility. Accordingly, content algorithms might encourage care only toward content (or, perhaps, the persons, ideas, or causes represented in content) which reaches a minimum threshold of visibility. In one respect, then, we encounter the same obstacle here as we do with media repertoires: you cannot gaze attentively at that which you cannot or do not see.

We also have to consider, though, what it means to gaze with attentive care at that content that does acquire visibility. This, again, is content that often expresses or evokes strong emotions. An example here, mentioned in chapters 1 and 4, is inspiration porn: objectifying content that shows disabled people doing fairly mundane things but uses framing language that portrays them as having done something bordering on miraculous. To reiterate a claim from chapter 1: the subjects of such content are not praised for (e.g.) having performed well in spite of the structural ableism that makes merely existing in the world as disabled more difficult than it is for nondisabled people. It celebrates their having done something *in spite of* their disability/disabilities. The framing is thus of disability, and not ableism, as the primary obstacle to accomplishment.

Attentive care, then, requires not subjecting people to narratives that they do not want to be party to; it is about not imposing narratives upon people, but letting them select for themselves what stories to tell. This becomes complicated when we think about persons with significant cognitive or intellectual disabilities who might lack the capacity to express their own narratives, or even to express their preferences about the narratives that others subject them to. Here, as at other points in this dissertation, Eva Kittay's reflections on parenting her (in her own words) significantly cognitively and physically disabled daughter are insightful. Kittay encourages readers to recognize the "fiction of independence and self-reliance"; she argues that we must recognize that some impairments "affect the capacity for self-determination" (2019, 149). In such cases, Kittay writes, care is about responsiveness to needs and to "legitimate wants" without diminishing or disparaging the cared-for (ibid, 137; 139). Genuine care requires that the one-caring take the "flourishing of the cared-for" as their ultimate value (ibid, 137). Thus, even when we cannot be sure of how another views themselves or how they wish to be viewed, we can look toward their needs and wants — however these may be expressed — and look to satisfy those without doing harm to the cared-for.

This logic can be extended to all of those who we don't know, or those whose ways of seeing themselves we may never know, but who we encounter narratives about online. When we encounter such narratives, we can scrutinize them by asking whether we have reason to think they portray the individual in a way they would legitimately want, or whether the narrative presents them in such a patronizing way that it could work against their flourishing.

Conclusion

"The best possible view of a person is that obtained by an attentive gaze that is truly loving, and this occurs when the view is not distorted by, for example, prejudice," write Carrasco and Alegría (2023 49). Coming to see justly is a moral accomplishment, which follows from the elimination of (or, at minimum, awareness of and conscious work against) prejudice. This holds both for the more traditional domains of morality — action — but also, importantly, for non-actionable ethics. This chapter has focused on attentive care as an ameliorative ethical disposition. It has looked at how we might, in cases like those described in earlier chapters, achieve the 'best possible view' that Carrasco and Alegría champion. As such, it is able to help address cases where our wrongs are not grounded in a particular action. Accordingly, so too can it help in cases where our obligations — what we should be doing instead — are not actionable.

Maintaining such an ethic is important for helping us clarify the nature of our affective obligations toward others, and our obligations toward their emotional displays. This is especially important to add to any ethic because, historically, ethics has either disregarded emotions entirely (for instance, a Kantian focus on doing the right thing out of a sense of duty regardless of how it *feels*) or has viewed them as beyond the scope of morality or moral critique. Social media in particular allows discourses like so-called therapy-speak, which render emotions beyond the bounds of moral critique, to proliferate.

One of the problems with an ethic of attentive care is that, by virtue of its amenability to cases where moral 'rightdoing' is not reflected in any observable action, it can be difficult to assess whether a just and loving gaze has actually been taken up. That is, what (if any) evidence do we have of how others see? Though I have been intentionally dispensing with cases where morality is solely action-based, I cannot deny that a just and loving gaze is often made obvious through people's actions. That is, I am arguing against an *exclusively* action-grounded ethic, not arguing that action has no proper place at all in morality.

Bibliography

Aaltola, E. (2014). Affective empathy as core moral agency: Psychopathy, autism and reason revisited. *Philosophical Explorations*, *17*(1), 76–92. https://doi.org/10.1080/13869795.2013.825004

Abramson, K. (with Harman, E.). (2024). *On gaslighting*. University Press.

Ahmed, S. (2003). The politics of fear in the making of worlds. *International Journal of Qualitative Studies in Education*, *16*(3), 377–398. https://doi.org/10.1080/0951839032000086745

Ahmed, S. (2004). Affective Economies. *Social Text*, *22*(2), 117–139. https://doi.org/10.1215/01642472-22-2_79-117

Ahmed, S. (2010). *The promise of happiness*. Duke University Press.

Ahmed, S. (2014). *The cultural politics of emotion* (Second edition). University Press.

Ahmed, S. (2017). *Living a feminist life*. Duke University Press.

Albrecht, G. L., & Devlieger, P. J. (1999). The disability paradox: High quality of life against all odds. *Social Science & Medicine (1982)*, *48*(8), 977–988. https://doi.org/10.1016/s0277-9536(98)00411-0

Alexander, M. (2023, June 7). *Selfish or self-care—Is this the end of therapy speak?* The Standard. https://www.standard.co.uk/lifestyle/therapy-speak-mental-health-tiktok-b1074680.html

American Psychological Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed). American psychiatric association.

Archer, A., & Matheson, B. (n.d.). Commemoration and Emotional Imperialism. *Journal of Applied Philosophy*, *n/a*(n/a). https://doi.org/10.1111/japp.12428

Archer, A., Mills, G., & University of Arkansas Press. (2019). Anger, Affective Injustice, and Emotion Regulation. *Philosophical Topics*, *47*(2), 75–94. https://doi.org/10.5840/philtopics201947216

Archer, A., & Robb, C. (2025). Ethics of Parasocial Relationships. In M. Betzler & J. Löschke (Eds.), *The Ethics of Relationships: Broadening the Scope* (p. 0). Oxford University Press. https://doi.org/10.1093/9780198928188.003.0012

Armah, E. (2022). *Emotional justice: A roadmap for racial healing* (First edition.). Berrett-Koehler Publishers, Inc.

Baier, A. (1994). *Moral prejudices: Essays on ethics*. Harvard University Press.

Barnes, E. (2016). *The minority body: A theory of disability* (First edition.). Oxford University Press.

Barnes, E. (2023). *Health problems: Philosophical puzzles about the nature of health*. Oxford University Press, Incorporated.

Bartky, S. L. (1990). *Femininity and domination: Studies in the phenomenology of oppression* (Thinking Gender). Routledge.

Beeker, T., Mills, C., Bhugra, D., te Meerman, S., Thoma, S., Heinze, M., & Peter, S. von. (2021). Psychiatrization of Society: A Conceptual Framework and Call for Transdisciplinary Research. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.645556

Bell, M. (2013). *Hard feelings: The moral psychology of contempt*. Oxford University Press.

Bell, M. (2023). The Wrong of Affective Dismissal and Its Place in an Account of Affective Injustice. *Philosophical Topics*, *51*(1), 239–264.

Benhabib, S. (1987). The Generalized Other and the Concrete Other: The Kohlberg-Gilligan Controversy and Feminist Theory. In D. Cornell & S. Benhabib (Eds.), *Feminism as critique: Essays on the politics of gender in late-capitalist societies*. Polity.

Benkler, Y. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (1st ed., pp. xii–xii). Yale University Press.

Bigelow, E. (2024). Anger, Affective Injustice, and Taylor Swift's "mad woman." In *Taylor Swift and Philosophy: Essays From the Tortured Philosophers Department*. John Wiley & Sons, Inc.

Blum, D. (2023, February 11). How the Language of Therapy Took Over Dating. *The New York Times*. https://www.nytimes.com/2023/02/11/style/therapy-speak-dating.html

Boleyn–Fitzgerald, P. (2003). Care and the Problem of Pity. *Bioethics*, *17*(1), 1–20. https://doi.org/10.1111/1467-8519.00318

Bollen, C. (2023). Towards a Clear and Fair Conceptualization of Empathy. *Social Epistemology*, *37*(5), 637–655. https://doi.org/10.1080/02691728.2023.2227963

Borgatti, S. P., & Foster, P. C. (2003). The Network Paradigm in Organizational Research: A Review and Typology. *Journal of Management*, *29*(6), 991–1013. https://doi.org/10.1016/S0149-2063(03)00087-4

Bowden, P. (1998). Ethical Attention: Accumulating Understandings. *European Journal of Philosophy*, *6*(1), 59–77. https://doi.org/10.1111/1468-0378.00050

boyd, danah m., & Ellison, N. B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210–230. https://doi.org/10.1111/j.1083-6101.2007.00393.x

Brown, É. (2018). Propaganda, Misinformation, and the Epistemic Value of Democracy. *Critical Review*, *30*(3–4), 194–218. https://doi.org/10.1080/08913811.2018.1575007

Buolamwini, J. (2023). *Unmasking AI: My Mission to Protect What Is Human in a World of Machines* (1st ed.). Random House Publishing Group.

Bynum, B. (2000). Discarded Diagnoses. *The Lancet*, *356*(9241), 1615. https://doi.org/10.1016/S0140-6736(05)74468-8

Campbell, S. M., & Stramondo, J. A. (2023). Expressed Ableism. *Ergo an Open Access Journal of Philosophy*, *9*(0), Article 0. https://doi.org/10.3998/ergo.3585

Carman, M. (2023). Harmony, Disruption, and Affective Injustice: Metz and the Capacity for Harmonious Relationship. *Ethical Theory and Moral Practice*. https://doi.org/10.1007/s10677-022-10362-0

Carr, C. T., & Hayes, R. A. (2015). Social Media: Defining, Developing, and Divining. *Atlantic Journal of Communication*, *23*(1), 46–65. https://doi.org/10.1080/15456870.2015.972282

Carrasco, M. A., & Alegría, D. (2023). "Loving Attention" in Murdoch, Smith, and the Ethics of Care. *CR: The New Centennial Review*, *23*(3), 45–69.

Chapman, R. (2019). Autism as a Form of Life: Wittgenstein and the Psychological Coherence of Autism. *Metaphilosophy*, *50*(4), 421–440. https://doi.org/10.1111/meta.12366

Chapple, M., Davis, P., Billington, J., Williams, S., & Corcoran, R. (2022). Challenging Empathic Deficit Models of Autism Through Responses to Serious Literature. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.828603

Chemaly, S. (2018). *Rage becomes her: The power of women's anger* (First Atria books hardcover edition.). Atria Books.

Cheng, A. A. (2000). *The melancholy of race*. Oxford University Press.

Cherry, M. (2023). Affective Stereotype Threat as Affective Injustice. *Philosophical Topics*, *51*(1), 135–148.

Cherry, M. V. (2021). *The case for rage: Why anger is essential to anti-racist struggle*. Oxford University Press.

Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and Them: Intergroup Failures of Empathy. *Current Directions in Psychological Science*, *20*(3), 149–153. https://doi.org/10.1177/0963721411408713

Cinnamon, M. (2024). *White Paper: Interventions for Online Harassment of Journalists | Berkman Klein Center*. https://cyber.harvard.edu/publication/2024/white-paper-interventions-online-harassment-journalists

Clare, E. (2017). *Brilliant Imperfection: Grappling with Cure*. Duke University Press.

Cudd, A. E. (2006). *Analyzing oppression*. Oxford University Press.

Davey, M. (2003, August 29). Faith Healing Gone Wrong Claims Boy's Life. *The New York Times*. https://www.nytimes.com/2003/08/29/us/faith-healing-gone-wrong-claims-boy-s-life.html

Davis, E. (2016). Typecasts, Tokens, and Spokespersons: A Case for Credibility Excess as Testimonial Injustice. *Hypatia*, *31*(3), 485–501.

Davis, N. A. (2005). Invisible Disability. *Ethics*, *116*(1), 153–213. https://doi.org/10.1086/453151

Della Ratta, D. (2020). Digital Socialism Beyond the Digital Social: Confronting Communicative Capitalism with Ethics of Care. *TripleC (Cognition, Communication, Co-Operation): Open Access Journal for a Global Sustainable Information Society*, *18*(1), 101–115. https://doi.org/10.31269/triplec.v18i1.1145

Dotson, K. (2012). A Cautionary Tale: On Limiting Epistemic Oppression. *Frontiers: A Journal of Women Studies*, *33*(1), 24–47. https://doi.org/10.5250/fronjwomestud.33.1.0024

Downey, D. B., & Conroy, J. W. (with Thornburgh, D., & Thornburgh, G.). (2020). *Pennhurst and the struggle for disability rights*. The Pennsylvania State University Press.

DuFord, R. (2022). *Solidarity in Conflict: A Democratic Theory* (1st ed.). Stanford University Press. https://doi.org/10.1515/9781503630703

Dunne, G., & Kotsonis, A. (2024). Carving at the Joints: Distinguishing Epistemic Wrongs from Epistemic Harms in Epistemic Injustice Contexts. *Episteme*, 1–14. https://doi.org/10.1017/epi.2023.62

Eriksson Krutrök, M., & Lindgren, S. (2022). Social media amplification loops and false alarms: Towards a Sociotechnical understanding of misinformation during emergencies. *The Communication Review*, *25*(2), 81–95. https://doi.org/10.1080/10714421.2022.2035165

Espinosa, C. R. (2018). "They Stay until They Die." *Human Rights Watch*. https://www.hrw.org/report/2018/05/23/they-stay-until-they-die/lifetime-isolation-and-neglect-institutions-people

Falbo, A. (2022). Hermeneutical Injustice: Distortion and Conceptual Aptness. *Hypatia*, *37*(2), 343–363. https://doi.org/10.1017/hyp.2022.4

Fanon, F. (with Sartre, J.-P., Farrington, C., & Grove Press). (1968). *The wretched of the earth* (First Black Cat Edition.). Grove Press, Inc.

Fanon, F. (with Markmann, C. L.). (1982). *Black skin, white masks* (1st Evergreen ed.). Grove Press.

Fishbein, R. (Guest Expert), Limbong, A. (Host). (2023). *"Therapy speak" is everywhere, but it may make us less empathetic* [Broadcast]. NPR.

https://www.npr.org/2023/04/13/1169808361/therapy-speak-is-everywhere-but-it-may-make-us-less-empathetic

Fisher, B., & Tronto, J. C. (1990). Toward a Feminist Theory of Care. In E. K. Abel & M. K. Nelson (Eds.), *Circles of care: Work and identity in women's lives*. State University of New York Press.

Franks, A. M., & Caldwell, D. (2023). Unpacking the Use of Therapy-Speak in Scholarly Writing. *American Journal of Pharmaceutical Education*, *87*(3), ajpe9030. https://doi.org/10.5688/ajpe9030

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press.

Fritts, M. (2023). *Echo Chambers and Social Media: On the Possibility of a Tax Incentive Solution*.

Fritts, M., & Cabrera, F. (2022). Fake News and Epistemic Vice: Combating a Uniquely Noxious Market. *Journal of the American Philosophical Association*, *8*(3), 454–475. https://doi.org/10.1017/apa.2021.11

Frye, M. (1983). *The politics of reality: Essays in feminist theory* (First edition., The Crossing Press Feminist Series). Crossing Press.

Gallegos, F. (2022). Affective Injustice and Fundamental Affective Goods. *Journal of Social Philosophy*, *53*(2), 185–201. https://doi.org/10.1111/josp.12428

Garbes, A. (2022). *Essential labor: Mothering as social change* (First edition.). Harper Wave, an imprint of HarperCollins Publishers.

Gardner, M. (2021). What is Harming? In J. McMahan, T. Campbell, J. Goodrich, & K. Ramakrishnan (Eds.), *Principles and Persons: The Legacy of Derek Parfit* (pp. 381–395). Oxford University Press.

Gendron, C. (2016). Moral attention: A comparative philosophical study. *Journal of Moral Education*, *45*(4), 373–386. https://doi.org/10.1080/03057240.2016.1208079

Gerwig, G., Rudin, S., Bush, E., O'Neill, E., Ronan, S., Metcalf, L., Letts, T., Feldstein, B., Chalamet, T., Hedges, L., Smith, L., Henderson, S., Rush, O., Rodrigues, J., Scott, M., Levy, S., Houy, N., Brion, J., Napier, A., … A24. (2017). *Lady Bird* [Video recording]. A24.

Gillespie, T. (2020). Content moderation, AI, and the question of scale. *Big Data & Society*, *7*(2), 2053951720943234. https://doi.org/10.1177/2053951720943234

Gillespie, T. (2022). Do Not Recommend? Reduction as a Form of Content Moderation. *Social Media + Society*, *8*(3), 20563051221117552. https://doi.org/10.1177/20563051221117552

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Harvard University Press.

Goodley, D. (2014). *Dis/ability studies: Theorising disablism and ableism* (1st Edition.). Routledge Publisher. https://doi.org/10.4324/9780203366974

Goodley, D., Liddiard, K., & Runswick-Cole, K. (2018). Feeling disability: Theories of affect and critical disability studies. *Disability & Society*, *33*(2), 197–217. https://doi.org/10.1080/09687599.2017.1402752

Gottfried, J., & Park, E. (2025, November 20). Americans' Social Media Use 2025. *Pew Research Center*. https://www.pewresearch.org/internet/2025/11/20/americans-social-media-use-2025/

Guan, S.-S. A., Hain, S., Cabrera, J., & Rodarte, A. (2019). Social Media Use and Empathy: A Mini Meta-Analysis. *Social Networking*, *08*(04), 147–157. https://doi.org/10.4236/sn.2019.84010

Hasebrink, U., & Domeyer, H. (2012). *Media repertoires as patterns of behaviour and as meaningful practices: A multimethod approach to media use in converging media environments*. *9*(2).

Hasebrink, U., & Popp, J. (2006). *Media repertoires as a result of selective media use. A conceptual approach to the analysis of patterns of exposure*. *31*(3), 369–387. https://doi.org/10.1515/COMMUN.2006.023

Held, V. (2006). *The ethics of care: Personal, political, and global*. Oxford Univ. Press.

Hochschild, A. R. (1983). *The managed heart: Commercialization of human feeling*. University of California Press.

Hughes, B. (2019). Invalidating Emotions in the Non-Disabled Imaginary: Fear, Pity and Disgust. In N. Watson, A. Roulstone, & C. Thomas (Eds.), *Routledge Handbook of Disability Studies* (2nd ed., pp. 89–101). Routledge. https://doi.org/10.4324/9780429430817-8

Jaggar, A. M. (1989). Love and knowledge: Emotion in feminist epistemology. *Inquiry*, *32*(2), 151–176. https://doi.org/10.1080/00201748908602185

Kafer, A. (2013). *Feminist, Queer, Crip*. Indiana University Press. http://ebookcentral.proquest.com/lib/washington/detail.action?docID=1189107

Kane, G. C., Alavi, M., Labianca, G. (Joe), & Borgatti, S. P. (2014). What's Different About Social Media Networks? A Framework and Research Agenda. *MIS Quarterly*, *38*(1), 275–304.

Kane, L. W. (2020). Accountability and Community on the Internet: A Plea for Restorative Justice. *Journal of Applied Philosophy*, *37*(4), 594–611. https://doi.org/10.1111/japp.12427

Kapp, S. K., Gillespie-Lynch, K., Sherman, L. E., & Hutman, T. (2013). Deficit, difference, or both? Autism and neurodiversity. *Developmental Psychology*, *49*(1), 59–71. https://doi.org/10.1037/a0028353

Kecmanovic, D. (1983). Psychiatrization: A General View. *International Journal of Social Psychiatry*, *29*(4), 308–312. https://doi.org/10.1177/002076408302900411

Kennett, J. (2006). Do psychopaths really threaten moral rationalism? *Philosophical Explorations*, *9*(1), 69–82. https://doi.org/10.1080/13869790500492524

Kittay, E. F. (2019). The Completion of Care—The Normativity of Care. In E. F. Kittay (Ed.), *Learning from My Daughter: The Value and Care of Disabled Minds* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780190844608.003.0010

Koehn, D. (1998). *Rethinking Feminist Ethics: Care, trust and empathy* (1st ed.). Routledge. https://doi.org/10.4324/9780203015650

Kohlberg, L. (1981). *Essays on moral development* (1st ed.). Harper & Row.

Konrad, A. M. (2021). Access Fatigue: The Rhetorical Work of Disability in Everyday Life. *College English*, *83*(3), 179–200.

Kurth, C. (2022). Inappropriate emotions, marginalization, and feeling better. *Synthese*, *200*(2), 155. https://doi.org/10.1007/s11229-022-03619-9

Larson, S. (2021, May 18). It's Vital That Police Better Understand Autism Spectrum Disorder. *HealthCity*. https://healthcity.bmc.org/its-vital-police-better-understand-autism-spectrum-disorder/

Lavallee, Z., & Gagné-Julien, A.-M. (2024). Affective injustice, sanism and psychiatry. *Synthese*, *204*(3), 94. https://doi.org/10.1007/s11229-024-04731-8

Lewis, A. (2020, January 23). *Being Your Selves: Identity R&D on alt Twitter*. Ribbonfarm. https://ribbonfarm.com/2020/01/23/being-your-selves-identity-rd-on-alt-twitter/

Liebow, N. (2016). Internalized Oppression and Its Varied Moral Harms: Self-Perceptions of Reduced Agency and Criminality. *Hypatia*, *31*(4), 713–729.

Lockwood, P. L., Ang, Y.-S., Husain, M., & Crockett, M. J. (2017). Individual differences in empathy are associated with apathy-motivation. *Scientific Reports*, *7*(1), 17293. https://doi.org/10.1038/s41598-017-17415-w

Lorde, A. (1984). *Sister outsider: Essays and speeches*. Crossing Press.

Lugones, M. (2003). *Pilgrimages = Peregrinajes: Theorizing coalition against multiple oppressions* (Feminist Constructions). Rowman & Littlefield.

Luthar, B., & Črnič, T. O. (2017). Media repertoires and discursive communities: Studying audiences in the multimedia age. *Communications: The European Journal of Communication Research*, *42*(4), 415–439. https://doi.org/10.1515/commun-2017-0028

Luzsa, R., & Mayr, S. (2019). Links Between Users' Online Social Network Homogeneity, Ambiguity Tolerance, and Estimated Public Support for Own Opinions. *Cyberpsychology, Behavior, and Social Networking*, *22*(5), 325–329. https://doi.org/10.1089/cyber.2018.0550

Maiese, M. (2024). Problems for enactive psychiatry? Mindshaping, social normativity, and neurodiversity. *Philosophical Psychology*, 1–26. https://doi.org/10.1080/09515089.2024.2446631

Marinucci, M., Pancani, L., Aureli, N., & Riva, P. (2022). Online social connections as surrogates of face-to-face interactions: A longitudinal study under Covid-19 isolation. *Computers in Human Behavior*, *128*, 107102. https://doi.org/10.1016/j.chb.2021.107102

Matassi, M., Mitchelstein, E., & Boczkowski, P. (2022). Social media repertoires: Social structure and platform use. *The Information Society*, *38*(2), 133–146. https://doi.org/10.1080/01972243.2022.2028208

Metzl, J. (2009). *The protest psychosis: How schizophrenia became a Black disease*. Beacon Press.

Miller, P., Parker, S., & Gillinson, S. (2004). *Disablism: How to Tackle the Last Prejudice*. Demos.

Morris, A. (2016). 'You Just Feel Like Nothing': California to Pay Sterilization Victims—The New York Times. *The New York Times*.

Murdoch, I. (1997). *Existentialists and mystics: Writings on philosophy and literature* (P. J. Conradi, Ed.). Chatto & Windus.

Murray, P. (2006). Being in School? Exclusion and the Denial of Psychological Relaity. In D. Goodley & R. Lawthom (Eds.), *Disability and psychology: Critical introductions and reflections* (pp. 34–41). Palgrave Macmillan.

Nario-Redmond, M. (2019). *Ableism: The Causes and Consequences of Disability Prejudice* (1st edition). Wiley-Blackwell.

Nario-Redmond, M. R., Kemerling, A. A., & Silverman, A. (2019). Hostile, Benevolent, and Ambivalent Ableism: Contemporary Manifestations. *Journal of Social Issues*, *75*(3), 726–756. https://doi.org/10.1111/josi.12337

Ngai, S. (2004). *Ugly feelings*. Harvard University Press.

Nguyen, C. T. (2020). ECHO CHAMBERS AND EPISTEMIC BUBBLES. *Episteme*, *17*(2), 141–161. https://doi.org/10.1017/epi.2018.32

Nguyen, C. T. (2024). VALUE CAPTURE. *Journal of Ethics & Social Philosophy*, *27*(3), 469–505. https://doi.org/10.26556/jesp.v27i3.3048

Nicholson, L. J. (1983). Women, Morality, and History. *Social Research*, *50*(3), 514–536.

Noble, S. U. (2013). Google Search: Hyper-visibility as a Means of Rendering Black Women and Girls Invisible. *InVisible Culture*, (19). https://doi.org/10.47761/494a02f6.50883fff

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. University Press.

Noddings, N. (1984). *Caring, a feminine approach to ethics & moral education*. University of California Press.

Noddings, N. (2002). *Starting at home: Caring and social policy* (1st ed.). University of California Press. https://doi.org/10.1525/9780520927568

Noddings, N. (2010). Complexity in Caring and Empathy. *Abstraca*, (V), 6–12.

Noddings, N. (2013). *Caring: A relational approach to ethics & moral education* (2nd ed., 2, Second edition, updated.). University of California Press. https://doi.org/10.1525/j.ctt7zw1nb

Noll, S. (2018). Institutions for People with Disabilities in North America. In *The Oxford Handbook of Disability History*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190234959.013.19

Norlock, K. J. (2017). Online Shaming. *Social Philosophy Today*, *33*, 187–197. https://doi.org/10.5840/socphiltoday201762343

NUSSBAUM, M. (1985). “FINELY AWARE AND RICHLY RESPONSIBLE”: MORAL ATTENTION AND THE MORAL TASK OF LITERATURE. *Journal of Philosophy*, *82*, 516–529. (PHL1135773).

Obar, J. A., & Wildman, S. (2015). Social media definition and the governance challenge: An introduction to the special issue. *Telecommunications Policy*, *39*(9), 745–750. https://doi.org/10.1016/j.telpol.2015.07.014

Olkin, R. (2002). Could you hold the door for me? Including disability in diversity. *Cultural Diversity and Ethnic Minority Psychology*, *8*(2), 130–137. (2002-02415-005). https://doi.org/10.1037/1099-9809.8.2.130

Phillips, W., & Milner, R. M. (2017). *The ambivalent Internet: Mischief, oddity, and antagonism online*. Polity.

Pismenny, A., Pismenny, A., Eickers, G., & Prinz, J. (2024). Emotional Injustice. *Ergo an Open Access Journal of Philosophy*, *11*(0). https://doi.org/10.3998/ergo.5711

Pohlhaus, G. (2012). Relational Knowing and Epistemic Injustice: Toward a Theory of "Willful Hermeneutical Ignorance." *Hypatia*, *27*(4), 715–735.

Pohlhaus Jr., G. (2014). Discerning the Primary Epistemic Harm in Cases of Testimonial Injustice: Social Epistemology: A Journal of Knowledge, Culture, and Policy. *Social Epistemology: A Journal of Knowledge, Culture, and Policy*, *28*(2), 99–114.

Price, M. (2011). *Mad at school: Rhetorics of mental disability and academic life*. University of Michigan Press.

Price, M. (2015). The Bodymind Problem and the Possibilities of Pain. *Hypatia*, *30*(1), 268–284.

Price, M. (2024). *Crip spacetime: Access, failure, and accountability in academic life*. Duke University Press.

Purves, D. (2019). Harming as making worse off. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, *176*(10), 2629–2656.

Puryear, C., & Vandello, J. A. (2019). Inflammatory Comments Elicit Less Outrage When Made in Anonymous Online Contexts. *Social Psychological and Personality Science*, *10*(7), 895–902. https://doi.org/10.1177/1948550618806350

Radden, J. (1985). *Madness and reason*. G. Allen & Unwin.

Reeve, D. (2006). Towards a Psychology of Disability: The Emotional Effects of Living in a Disabling Society. In D. Goodley & R. Lawthom (Eds.), *Disability and psychology: Critical introductions and reflections*. Palgrave Macmillan.

Reeve, D. (2019). Psycho-Emotional Disablism: The Missing Link? In N. Watson, A. Roulstone, & C. Thomas (Eds.), *Routledge Handbook of Disability Studies* (2nd ed., pp. 102–116). Routledge. https://doi.org/10.4324/9780429430817-9

Reynolds, J. M. (2017). “I’d rather be dead than disabled”—The ableist conflation and the meanings of disability. *Review of Communication*, *17*(3), 149–163. https://doi.org/10.1080/15358593.2017.1331255

Reynolds, J. M. (2022). *The life worth living: Disability, pain, and morality*. University of Minnesota Press.

Roberts, J. (2021). Empathy Cultivation through (Pro)Social Media: A Counter to Compassion Fatigue. *Journalism and Media*, *2*(4), 819–829. https://doi.org/10.3390/journalmedia2040047

Robinson, F. (1999). *Globalizing care: Ethics, feminist theory, and international relations*. Westview Press.

Ruddick, S. (1989). *Maternal thinking: Toward a politics of peace*. Beacon Press.

Samuels, E. (2003). MY BODY, MY CLOSET: Invisible Disability and the Limits of Coming-Out Discourse. *GLQ: A Journal of Lesbian and Gay Studies*, *9*(1–2), 233–255. https://doi.org/10.1215/10642684-9-1-2-233

Saul, J. (2024). *Dogwhistles and Figleaves: How Manipulative Language Spreads Racism and Falsehood*. Oxford University Press.

Schalk, S. D. (2022). *Black disability politics*. Duke University Press.

Schieferdecker, D., & Wessler, H. (2017). Bridging Segregation Via Media Exposure? Ingroup Identification, Outgroup Distance, and Low Direct Contact Reduce Outgroup Appearance in Media Repertoires. *Journal of Communication*, *67*(6), 993–1014. https://doi.org/10.1111/jcom.12338

Schulman, S. (2016). *Conflict is not abuse: Overstating harm, community responsibility, and the duty of repair*. Arsenal Pulp Press.

Shapiro, J. P. (1993). *No pity: People with disabilities forging a new civil rights movement* (First edition.). Times Books.

Slaby, J. (2023). Structural Apathy, Affective Injustice, and the Ecological Crisis. *Philosophical Topics*, *51*(1), 63–83. (EP178302766). https://doi.org/10.5840/philtopics20235114

Søe, S. O. (2021). A unified account of information, misinformation, and disinformation. *Synthese*, *198*(6), 5929–5949.

Solomon, R. C., & Stone, L. D. (2002). On “Positive” and “Negative” Emotions. *Journal for the Theory of Social Behaviour*, *32*(4), 417–435. https://doi.org/10.1111/1468-5914.00196

Sontag, S. (2002, December 1). Looking at War. *The New Yorker*. https://www.newyorker.com/magazine/2002/12/09/looking-at-war

Spelman, E. (1989). Anger and Insubordination. In A. Garry & M. Pearsall (Eds.), *Women, knowledge, and reality: Explorations in feminist philosophy*. Unwin Hyman.

Srinivasan, A. (2018). The Aptness of Anger. *Journal of Political Philosophy*, *26*(2), 123–144. https://doi.org/10.1111/jopp.12130

Stern, R., & Brackett, M. (2023, January 26). Advice | Traumatized? Gaslit? How to know if you’re misusing therapy words. *The Washington Post*. https://www.washingtonpost.com/wellness/2023/01/10/misuse-gaslighting-trauma-narcissism/

Stockdale, K. (2023). (Why) Do We Need a Theory of Affective Injustice? *Philosophical Topics*, *51*(1), 113–134. https://doi.org/10.5840/philtopics20235116

Stramondo, J. A. (2010). How an Ideology of Pity Is a Social Harm to People with Disabilities. *Social Philosophy Today*, *26*, 121–134. https://doi.org/10.5840/socphiltoday20102610

Svendsen, L. (2008). *A Philosophy of Fear*. Reaktion Books.

TED. (2014, June 9). *I’m not your inspiration, thank you very much | Stella Young* [Video recording]. https://www.youtube.com/watch?v=8K9Gg164Bsw

Tessman, L. (2023). An Apology for Inapt Emotions. *Philosophical Topics*, *51*(1), 189–220.

Tewes, C. (2024). Reconsidering the Double Empathy Problem: A source of exclusion and stigmatisation of autistic persons. *Phenomenology and Mind*, (27), 126–137.

Thaler, R. H., & Sunstein, C. R. (2003). Libertarian Paternalism. *American Economic Review*, *93*(2), 175–179. https://doi.org/10.1257/000282803321947001

Thomas, C. (1999). *Female forms: Experiencing and understanding disability*. Open University Press.

Thomas, E. F., Cary, N., Smith, L. G., Spears, R., & McGarty, C. (2018). The role of social media in shaping solidarity and compassion fade: How the death of a child turned apathy into action but distress took it away. *New Media & Society*, *20*(10), 3778–3798. https://doi.org/10.1177/1461444818760819

Thomas, E. F., McGarty, C., & Mavor, K. I. (2009). Transforming "Apathy Into Movement": The Role of Prosocial Emotions in Motivating Action for Social Change. *Personality and Social Psychology Review*, *13*(4), 310–333. https://doi.org/10.1177/1088868309343290

Thomson, R. G. (2009). *Staring: How we look*. Oxford University Press.

Tronto, J. C. (1993). *Moral boundaries: A political argument for an ethic of care*. Routledge.

Turkle, S. (1996). Virtuality and Its Discontents: Searching for Community in Cyberspace. In *The American prospect* (Issue 24, pp. 50–57).

Unruh, C. F. (2023). A Hybrid Account of Harm. *Australasian Journal of Philosophy*, *101*(4), 890–903. https://doi.org/10.1080/00048402.2022.2048401

Västfjäll, D., Slovic, P., Mayorga, M., & Peters, E. (2014). Compassion Fade: Affect and Charity Are Greatest for a Single Child in Need. *PLoS ONE*, *9*(6), e100115. https://doi.org/10.1371/journal.pone.0100115

Vishkin, A., & Tamir, M. (2023). Emotion Norms Are Unique. *Affective Science*, *4*(3), 453–457. https://doi.org/10.1007/s42761-023-00188-z

Watermeyer, B. (2017). "I Don't Have Time for an Emotional Life": Marginalization, Dependency and Melancholic Suspension in Disability. *Culture, Medicine, and Psychiatry*, *41*(1), 142–160. https://doi.org/10.1007/s11013-016-9503-x

Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. The MIT Press.

Weil, S. (with Kirkpatrick, D. K.). (1952). *The Need for Roots: Prelude to a Declaration of Obligations Towards the Human Being* (R. Schwartz, Trans.). Penguin Classics.

Wendell, S. (1996). *The rejected body: Feminist philosophical reflections on disability*. Routledge.

West, R., & Johnson, B. (2024). "I hope you aren't becoming woke": New politicised contours of online ableism in response to disability advocacy by 2022 Australian of the Year Dylan Alcott. *Australian Journal of Social Issues*, *59*(2), 517–531. https://doi.org/10.1002/ajs4.305

Whitney, S. (2018a). Affective Intentionality and Affective Injustice: Merleau-Ponty and Fanon on the Body Schema as a Theory of Affect. *The Southern Journal of Philosophy*, *56*(4), 488–515. https://doi.org/10.1111/sjp.12307

Whitney, S. (2018b). Affective Intentionality and Affective Injustice: Merleau-Ponty and Fanon on the Body Schema as a Theory of Affect. *Southern Journal of Philosophy*, *56*(4), 488–515. https://doi.org/10.1111/sjp.12307

Whitney, S. (2024). *Anger Gaslighting and Affective Injustice*.

Yap, A. S. (2017). Credibility Excess and the Social Imaginary in Cases of Sexual Assault. *Feminist Philosophy Quarterly*, *3*(4), Article 4. https://doi.org/10.5206/fpq/2017.4.1

Yergeau, M. (2018). *Authoring Autism: On Rhetoric and Neurological Queerness* (1st ed.). Duke University Press. https://doi.org/10.1515/9780822372189

Young, I. M. (2011). *Justice and the politics of difference* (Paperback reissue / with a new foreword by Danielle Allen.). Princeton University Press.

Young, I. M., & Nussbaum, M. (2011). *Responsibility for Justice*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195392388.001.0001