

## The Snares of Self-Hatred

Please refer to published version in *The Moral Psychology of Hate*, Rowman and Littlefield (2022).

### I. Introduction

In imagining self-hatred, we see a person turned in on herself. Her suffering seems driven by two seemingly antagonistic, but nonetheless complementary forces. From one angle we see a repeated pattern of self-directed hostility. From another we see, curiously, a *receptivity* to that hostility. Unimpeded, these movements together form a cyclical and dynamic progression difficult to resolve or subdue.

As with certain other self-reflexive emotions, such as guilt and shame, our understanding of self-hatred may be aided by views of the mind which posit an internalized other whose perspective on oneself embodies and focuses a set of concerns and values, and whose perspective one is in some sense *vulnerable* to. To feel guilt for some transgression is not solely to feel the anger that one would feel toward another's trespasses, now directed back onto oneself as an object of that anger; it is at the same time to react to that anger—perhaps, for example, to accept it as deserved, or to welcome the lashing of one's bad conscience. To feel shame before oneself is not just to see oneself in some compromising way, it is to *feel compromised* by one's own gaze.

Likewise, the person who hates herself does not feel the hatred that she might have for another, simply taking herself as object of that attitude. She is not merely the seat of an internalized hostile voice and perspective that she may, for example, react to with indifference. She does not only tell herself that she is “worthless” but will typically *feel* herself so in response. And her suffering may not just result from pain she is inclined to *inflict*, but suffering that, in some sense, she is inclined *to suffer*. But how is this so? How, in self-hatred, does one become not only subject to, but *vulnerable* and even *receptive* to one's own hostility?

### II. Authority and Vulnerability

Begin with an account of self-hatred that has been inherited at least by our culture, if not by contemporary philosophers and psychologists, and which finds its sources in Freud. That self-

hatred can impede one's hopes for happiness is obvious. What is less obvious, according to Freud, is that it is the price we pay for developing a superego, and thereby a moral conscience, at all. On this picture, the self-hating man is relentlessly cruel and abusive to himself, seeks out forms of self-punishment and self-sabotage, castigates himself for actions he has not in fact performed, and may even, as a "pale criminal," act in ways that affirm the persecutory stance of his own psyche (Freud 1916). This is because he is, though unaware of it, wracked by feelings of guilt—guilt which is, in part, a response to motives and impulses that he may be able to conceal from others but which, upon developing a superego, he cannot conceal from himself.

But Freud also sought to explain why one's hostility toward oneself will tend to be disproportionate, or seemingly wholly indifferent to whether one has in fact done, intended, or even merely entertained the thought of doing something that would warrant feelings of guilt. This is because, he suggests, the superego does not only serve as a seat of moral conscience, "keeping a watch over the actions and intentions of the ego and judging them... exercising a censorship" (Freud 1989, 136). It also redirects the energy of a set of hostile, aggressive, and destructive impulses part of our shared natural endowment. One's mind must find release of this aggressive energy somewhere, and so, this explanation goes, one's superego redirects it back onto oneself, channeling this aggression through pangs of bad conscience rather than onto others whom one must live, relatively peacefully with, in society.<sup>1</sup> Thus, we arrive at a proposal that promises to explain not only how and why self-hatred arises, but why it may be so ubiquitous, and why it can be so persistent, as well.

But suppose we take the seeming sources of self-hatred as genuine. What people seem to hate themselves for, again, does not seem limited to those things that would give rise to the feeling of guilt. This will often include nonmoral characteristics of one's self, including much that one could not be, in any sound sense, morally responsible for. Of course, one might insist that guilt is pernicious in just this way: it captures and holds in distortive construal a wide range of things

---

<sup>1</sup> As Freud writes, "Conscience arises through the suppression of an aggressive impulse, and . . . it is subsequently reinforced by suppressions of the same kind" (Freud 1989, 92). John Deigh (1996) argues that Freud's account of the emergence of the superego in *Civilization and Its Discontents* is an alternative to his earlier account, according to which the superego develops as the resolution of the Oedipal complex. My aims in this paper are not exegetical, and I will largely assume Deigh's reading of Freud here. In *Mourning and Melancholia* (1915), Freud characterizes the self-hatred that is experienced by the "melancholic," which is also constituted by a person's redirection of hostility toward the lost loved object, onto oneself, once that object has been internalized, and is also expressed as moral castigation of the self. The self-hatred I will explore here differs from this form, as well, for reasons similar to the ones I will focus on here.

that one would not, if only one were perhaps more clearheaded or well-disposed, feel guilty about. Survivor's guilt, and the guilt commonly experienced by victims of sexual assault, for example, may testify to this. But we might also consider that self-hatred, after all, may take different forms—and that at least one form would itself be distorted by thinking of it as a manifestation of guilt, even confusedly or pathologically experienced. We might also consider that while self-hatred may be a condition we become liable to upon becoming socialized, the social conditions that give rise to it and influence its content are more contingent than Freud allows. It is not just the constraints of society, at all, but the particular and concrete arrangements of a particular and concrete social reality that will influence and inform the manifestations of self-hatred that arise within it.

To illustrate the kind of self-hatred I have in mind, consider James Baldwin's 1964 novel, *Another Country*.<sup>2</sup> At the center of this novel is the unraveling and eventual self-destruction of Rufus Scott—a young Black man living in New York in the late 1950s. Once self-assured and full of life, we are introduced to a man who is a ghost of his former self. It is a lingering mystery to Rufus and his friends and family what has happened. While it is suggested, after his death, that Rufus was wracked with guilt—guilt felt in response to his brutal and violent treatment of his lover, Leona, a white woman from Georgia—this explanation, offered by a character who admits that he never cared much for Rufus anyway, strikes one as both unimaginative and incomplete. There seems to have been some other force at work here, some other explanation for Rufus's self-hatred: one which explains his subsequent mistreatment of Leona and his subsequent guilt for that mistreatment.

There are two other aspects of Rufus's self-hatred worth pausing on. The first is the quality of Rufus's suffering. In response to his hostility he feels diminished, hopeless, and alone. He

---

<sup>2</sup> I use this example to fix our attention. There are many different forms of what we could rightly call "self-hatred," perhaps understood most abstractly as some kind of persistent negative stance taken toward oneself, given some view that one has of oneself. It is not an ambition of mine to suggest that the kind of self-hatred under investigation here is the one true kind, or the only kind: just one genuine kind with ethical and political significance. I also proceed on the assumption that certain writers and artists, including Baldwin, successfully describe genuine human experiences which give rise to certain philosophical questions, and that Freud, as well as other psychoanalytic thinkers, are, as Richard Moran puts it, adding to the forms of understanding of our minds and others at the level of "folk psychology": "the hopes and fears, pains and experiences we relate to each other in daily life, and not states or processes defined either neurologically or computationally" (Moran 2001, 7). There are no doubt neurological explanations for aspects of hatred, as Berit Brogaard has recently explored, for example, the tendency for hatred to become all-consuming "correlates with an increase in norepinephrine and dopamine signaling in the brain's prefrontal cortex" which is the same thing that happens after "a big hit of methamphetamine or cocaine" (Brogaard 2020, 38). But this is not the level of explanation I am offering here.

experiences a kind of psychic pain and anguish that he comes to believe will only be alleviated should he no longer exist at all. The second aspect is that Rufus's self-hatred seems to perpetuate itself, and persist in spite of the love of Leona, as well as of his close friend, Vivaldo. It is repeatedly stressed that he loves them both, and they love him. But their reassurances are drowned out by his self-loathing: the hostility he feels toward himself is too persistent, or powerful, to be lessened or diminished by their love. Their attempts to express their love for him only seem to exacerbate his hostility—both toward himself and toward them.

These two features of Rufus's suffering inform my focus and investigation here. But again, my central question, roughly stated, is how self-hatred manifests not only in hostility toward oneself, but in a certain kind of vulnerability to that hostility: how someone becomes, as Rufus becomes, susceptible to one's own venom. Because it isn't obvious that these two components of self-hatred could be understood as operating independently of one another, I will frame this investigation in light of another question that has arisen from contemporary philosophical engagement with Freud's views—a question that gives rise to what I will refer to as the *problem of authority*. The problem of authority arises given a Kantian approach to Freudian psychology and a Kantian ambition of explaining how the superego, as the purported seat of moral conscience. The problem is to plausibly explain how the superego gains *moral authority* over the ego. Beginning with this question will, I hope, serve as a way of seeing why the question I pose about self-hatred is an open one.

The contemporary philosophers puzzled by the problem of authority have remarked on several striking parallels between Kant's view of moral motivation and Freud's discussion of the superego.<sup>3</sup> And Freud himself suggests that the superego just is the Kantian "moral law within us." For Kantian moral psychology, this is of significance because Freud's account of the superego seems to be a promising and perhaps naturalized account of the Kantian conception of moral conscience, and a Kantian conception of guilt. The superego bears the strictness and disciplinary qualities of this ideal. As a part of a more general moral-psychological view, Freud's proposal also seems to capture the idea that to violate a moral law one acts in disobedience to some governing authority within oneself, and typically because of some motive of self-gratification. And, importantly, in developing a superego, a person becomes disposed to subject himself to feelings of bad conscience which arise independently of concerns about having his

---

<sup>3</sup> Velleman (1999); Deigh (1996, 1999); Scheffler (1992); Longuenesse (2012).

transgressions witnessed or punished by actual others—a necessary condition for possessing a recognizably moral conscience, at all.

But the problem of authority arises once we see that mere internalization of an aggressive figure cannot adequately account for the disposition to feel guilt should we understand this emotion as also, necessarily, a response to a perceived violation of one's moral obligations, rather than as a kind of anxiety one feels in response to being punished for that violation. In a canonical, early piece of analytic-philosophical engagement with Freudian psychology, David H. Jones (1966) argues that Freud does not adequately distinguish between refraining from violating a moral duty because one fears the hostility of one's superego and refraining because one fears an external threat from an actual other. Thus, Jones concludes, Freud's views cannot accommodate genuine moral motivation nor a proper conception of guilt.

The problem, David Velleman argues, is that Freud does not give a plausible account of how the ego views the superego so that it is seen as a source of morally authoritative demands and punishments, rather than as a source of demands the subject might simply take a dismissive or defiant attitude toward, and threats he simply fears, “despite issuing from a part of himself” (Velleman 1999, 544, n. 35). We must account for how the ego, as Velleman puts it, “buys into” the superego's demands (1999, 563). According to Velleman's reconstruction of Freud, the ego “buys into” the superego's demands when it views the superego not just as an internalized source of demands and threats, but also as an ideal figure that it both loves and strives to emulate. But fear of the loss of the love of an internalized figure is, just like the fear of that figure's anger, not yet to grant that figure moral authority. So, furthermore, Velleman attributes to the ego a capacity for normative judgment, exercised when the subject reflects on what ideals she ought to have. While her ego-ideal is originally shaped around her parental figures, as she develops and matures as an autonomous moral agent she must determine for herself what the content of this ideal should really be, thereby coming to act under laws, or according to demands or ideals, that she has given to herself.

My concern here is not with whether the problem of authority is successfully met by Velleman's proposal—or indeed, even whether it should be.<sup>4</sup> One might think, after all, that our

---

<sup>4</sup> It is precisely the peculiarity of moral authority explicated along Kantian lines as either fully intelligible or necessary to recognizably moral life that Williams criticizes in *Shame and Necessity* (1993) and elsewhere.

moral emotions lie somewhere in between the crudest fears of being caught, punished, or exiled, and a response felt in violation of a law that one has given to oneself. Nonetheless, considering the form of the problem provides a path forward in considering what I can now label, in contrast to the problem of authority, the *problem of vulnerability*. First, as with the problem of authority, the problem of vulnerability arises in the course of attempting to understand a particular emotion—in this case, self-hatred—that, because of its self-reflexive nature, may be illuminated through the positing of an internalized other whose hostile perspective on oneself characterizes and partly constitutes the emotion in question. Second, as I’ll argue, the problem of vulnerability also arises once we see that the mere internalization of some hostile figure is not sufficient to account for self-hatred, given that one is also in some sense vulnerable to that figure’s hostile perspective. So, third, the solution will involve positing a way in which the ego views, or relates to, the internalized other, so that the subject experiences that hostility in a way that will account for the vulnerability one experiences in self-hatred. And the solution I will offer—like Velleman’s proposed solution to the problem of authority—will involve the ego viewing and experiencing the internalized other as an object of a certain kind of interpersonal *love*.

In order to see why mere internalization of a hostile figure is not sufficient to account for self-hatred, we must also see what kind of vulnerability we are attempting to account for. One might propose that if we can account for the vulnerability and receptivity to one’s own hostility that one experiences in guilt, we will, *a fortiori*, be able to account for these features of self-hatred—assuming, as Freud proposes, that self-hatred just is a manifestation of guilt. But as I’ve suggested, this proposal seems incomplete. Most notably, self-hatred can stem not from a concern about what one has done or is inclined to do, but rather, in response to qualities or traits that one believes oneself to possess. And while one might hate oneself for traits one believes best explain why one has acted (or is disposed to act) in morally wrong ways, feeling what Jean Hampton (1990) calls “moral hatred” for oneself, it is just as familiar that one can hate oneself because of one’s looks, one’s lack of talent, or membership in a class or group of people—and not because one implicitly believes that one has a duty or obligation not to possess these qualities, or that one is morally responsible for having such qualities in the first place. The sources of self-hatred are wide in kind, particular to the person, and not reducible to considerations of what one perceives oneself obligated to eliminate or alter. In these respects, the sources of self-hatred resemble the sources of shame.

A further distinction between the vulnerability one experiences in guilt and the vulnerability that one experiences in self-hatred can be drawn if we were also to assume that in order to feel guilt of the kind that motivates the problem of authority, one must have acted in violation of a law that one has established for oneself autonomously. But self-hatred is often notably, and for some, maddeningly, heteronomous. Imagine, for example, a woman who has spent her life understanding and rejecting traditionally feminine norms of behavior and appearance for good reasons, but who nonetheless begins to hate herself for her thinning hair and sagging skin.<sup>5</sup>

In coming to hate oneself, then, one does not need to have granted any kind of authority to one's internalized other. However, one nonetheless views that other as possessing forms of *epistemic*, *definitional*, and *evaluative* credibility about one's self. Imagine, for example, a man who has come to hate himself for his lack of ambition. First, in hating himself, he accepts the descriptive aspect of this self-assessment. Second, he does not simply hate his lack of ambition, understood as just one of his traits among others. He hates *himself* for his lack of ambition. And so, he accepts too, that this quality defines him: that his lack of ambition is of central significance to who he is. Hatred of oneself mirrors hatred of others in this respect: it tends to both globalize certain traits as defining the hated person and presents those traits as essential to the hated person's identity (see, for example, Fischer et al. 2018). And third, along with accepting this defining assessment of himself, he does not just view his lack of ambition as something he can accept dispassionately, with equanimity, or good humor: he is wounded, his emotional response evidence that his being an unambitious person has a certain kind of evaluative significance for him.

But what kind of "evaluative significance" is this? One might suggest that it stems from the general vulnerability that we experience in light of the perceptions and judgments of other human beings—a respect in which we are not vulnerable to, for example, natural disasters or disease, even when such things frustrate our ends or ravage our bodies. As discussed by P. F.

---

<sup>5</sup> Brogaard also discusses how certain psychic divisions are necessary to explain certain self-reflexive attitudes, such as self-blame. On her view, our conscience "plays the role of the cynical unfeeling judge, and the part of us that feels bad plays the role of the accused," and "the division of labor is one between reason and passion" (Brogaard 2020, 226). It's not obvious why the accusing role is identified with "reason" here, though perhaps an example she discusses later may help. She considers the self-hatred of women who (rationally) *endorse* certain misogynistic standards about women, and who then hate themselves when they fail to live up to those standards. But my point is that no such endorsement is necessary—a woman can hate herself while rejecting these standards.

Strawson (1973), in being wronged by another, what we object to is not solely, or sometimes centrally, the material injury or harm that their action has brought about: we object to what that action suggests about how we are being viewed by the wrongdoer, and what we believe, or fear, it suggests about ourselves—hence our emotional reaction of resentment when we can presume that they, given what they have done, regard us without due respect or good will. As Jeffrie Murphy has put this point,

Most of us tend to care about what others (at least some others, some significant group whose good opinion we value) think about us—how much they think we matter. Our self-respect is *social* in at least this sense, and it is simply part of the human condition that we are weak and vulnerable in these ways. (Murphy and Hampton 1990, 25.)

But the weakness and vulnerability one experiences in self-hatred is in one respect narrower, and in another respect broader, than the kind that Strawson identifies and that Murphy and others have elaborated on. First, we are not in fact vulnerable to the views of any hostile human being. There are actual human beings who may be hostile toward us, who are full of ill will and who view us in negative ways, and who may express their hostility in cruel and aggressive manners—but to whom we may react with dispassionate reciprocated hostility, management, or even indifference. Consider, for example, Bernard Williams’s discussion of Ajax’s shame and suicide after his delusional slaughter of a flock of sheep:

[Ajax] could not go on living... It was in virtue of the relations between what he expected of the world and what the world expects of a man who expects that of it. “The world” there is represented in him by an internalized other, and it is *not merely any other*; he would be as unimpressed by the contempt of some people as he would be led by the reassurances of others. (1993, 84–5, my emphasis)

Given that we are not emotionally vulnerable to the hostility of *any* other human perspective on ourselves, the mere internalization of a hostile perspective is not enough to secure the kind of vulnerability to ourselves we experience in self-hatred. I might react to this internalized perspective just as I react to actual others whose views I am, as Williams puts it, “unimpressed”

by. And we can spell out this lack of vulnerability along the lines I described earlier: I might not grant this voice epistemic credibility, disbelieving or easily dismissing what it says—for example, that I am a terrible singer, that I am overly cynical, or that I am too eager to please—as simply inaccurate. Or I might feel hatred for these qualities, but experience this in a localized, rather than globalized manner—hating some aspect of myself, without hating myself in light of that aspect, because I do not accept that it is something which defines who I am. Or I may even accept that I possess a certain trait or quality and that it defines me, but not experience this as having the kind of evaluative significance that seems necessary to experience both sides of self-hatred. It may not matter, in the relevant kind of way, for me to feel wounded or diminished by this self-assessment, even though I see that it is true that I have this quality or trait, and that it defines me.

Second, while Strawson tends to emphasize attitudes that we can reasonably expect or demand from one another—a modicum of good will and respect, for example—our hopes for how some people view us will not be limited to what we can reasonably expect or demand from them. Just as we can be wounded, hurt, and diminished by how certain others see us without presupposing that they owe it to us to see us any differently, the self-hating person’s emotional vulnerability to his own hostility is not one that presupposes a self-directed demand that he be seen by himself in certain ways rather than others. When others fail to see us in ways that we may hope for or need, but which we do not think can be reasonably demanded, we may not feel *resentment* in response to this hostility, but again, diminishment and anguish.<sup>6</sup> A person who is able to convince himself that he doesn’t deserve this hostility, and is thereby able to emotionally bolster himself against his own criticisms and protest their credibility, is already shoring up a lack of vulnerability one may need to *mitigate* one’s self-loathing.<sup>7</sup>

Here, then, is a restatement of the puzzle of vulnerability. In self-hatred, how does one view one’s internalized other, so that one experiences this figure as having epistemic, definitional, and evaluative credibility, where that evaluative credibility is evinced by a person’s suffering from feelings of diminishment and anguish in response to that other’s hostility? When one “buys into” the hostile perspective that one has on oneself, one is not experiencing that

---

<sup>6</sup> Though, for a plausible argument that resentment is in fact a form of anger that understands itself as unjustified, see Ulrika Carlsson’s, “Tragedy and Resentment” 2018.

<sup>7</sup> For discussion of how protest against how one is being seen by another, grounded in a demand or expectation to be seen differently, can oneself defend against emotional vulnerability, see George Yancy’s (2006) discussion of anger.

perspective as simply the perspective of “any other,” as Williams points out. So how is one experiencing it?

As I suggested, the answer to the problem of vulnerability I will propose stems from one more similarity it shares with the problem of authority. Their answers both rely on another basic attitude to explain how the emotion under investigation develops: our love of others. On Freud’s picture of the emergence of guilt, it is our love of our parents (understood in its most rudimentary form as a love for them given their responsiveness to our physiological needs), paired with our hostility toward them, that results in the emotional ambivalence that must be resolved through the development of the superego and thereby the disposition to feel guilt. For Velleman’s Freud, the superego gains its authority over the ego partly through the love that one has for one’s parental figures, which explains why they become the ego’s first ego-ideals: a form of love that also develops from a recognition of their ability to respond to one’s physiological neediness, but which, Velleman argues, is fundamentally responsive to their humanity, in the Kantian sense.<sup>8</sup> Following this lead, my answer to the problem of vulnerability will begin by first considering how and in what respects we are vulnerable, in general, to our loved ones. And upon isolating a particular kind of vulnerability to our loved ones that I will discuss here—one that arises from a kind of neediness for them not reducible to our physiological neediness nor reducible to our needs for respect—we will be able to see *why* we are vulnerable to ourselves in self-hatred, as well as why we are vulnerable in the specific ways that I have described.

### III. Love and the Relational Self

Interpersonal love renders us emotionally vulnerable to others, as well as susceptible to certain forms of suffering. According to Harry Frankfurt (1999) to love another is to will her well-being for her sake, and to experience her well-being as an extension of one’s own; thus, one is liable to suffer when she suffers. As Martha Nussbaum (2001) emphasizes, to love another human being is to love a finite, vulnerable, and fundamentally mortal creature. Grief, then, we might think, is a necessary development of that love. As Monique Wonderly (2017) has recently

---

<sup>8</sup> “What the child experiences in being loved by his parents, and what he responds to in loving them, is their capacity to anticipate and provide for his needs, often at the expense of their own interests. And this capacity of the parents is nothing other than their practical reason, or practical good sense, by which their immediate self-gratification is subordinated to rational requirements” (Velleman 1999, 556).

discussed, developing the psychological views of John Bowlby, certain forms of love manifest themselves in attachments to our loved ones, and thus, we are left feeling less competent and confident in our agency when we are separated from them. And, of course, we tend to empathize, sympathize, and even sometimes emotionally identify with our loved ones; thus, their joys will be our joys, but so too, their pain and suffering will be ours as well.

Here I want to isolate one more way in which we are vulnerable to our loved ones. Begin with an observation by Herbert Morris. In a discussion of the various forms of suffering one may experience in feeling guilty, Morris discusses a case in which one wrongs not just a person who is a member of one's moral community, but who is also a person whom one loves. Alongside the pain of guilt and separation, and the pains of empathy, sympathy, emotional identification, Morris describes a distinctive kind of pain that one may feel in recognition that one is responsible for the dissolution of a union that one was once party to. As he writes,

To be cut off from what we love is intensely painful, and the pain involved in guilt resembles this. But there is more involved, for I have suggested that in union, best exemplified in love, there is an intensely satisfying *feeling of wholeness or completeness*... In cutting oneself off from others one comes to see oneself as being cut off, not whole, as if one had destroyed what one loved and thus also *destroyed a part of oneself*. (Morris 1971, 426, my emphasis.)

What explains this “intensely satisfying feeling of wholeness or completeness” in union? Why, in losing this union or the relationship necessary to sustain it, might one feel as though one has also “destroyed a part of oneself”?

These experiences can be captured and explained should we take seriously the idea that not only does our self-respect depend on the attitudes of at least some others, but, as philosophers who have argued that the self is essentially relational have discussed, our identities, too, are so vulnerable. The attitudes of our loved ones in particular play a central role in the formation and maintenance of our identities. Developmentally, it is within the context of intimate relationships that one receives the kind of interaction that will form one's most basic sense of self: one learns that one is an agent, and that one has needs that need to be met, and so on, through being recognized and responded to as an agent with such needs by one's early caregivers.<sup>9</sup> And it is

---

<sup>9</sup> For discussion of the “intersubjective self,” its reliance on relationships of mutual recognition with others, and its development out of the relationship between infant and mother, see Jessica Benjamin (1988).

also within intimate relationships that we continue to discover, develop, and maintain who we are: one's self-conceptions become richer, and more developed, in tandem with and in response to the conceptions that others have of oneself. One both becomes and discovers what kind of sister one is, for example, in relation to one's siblings, and as one develops in light of their conceptions of what kind of sister one is (see Lindemann 2014). Moreover, the kind of first-personal authority that we sometimes have in defining our self-conception (i.e., the qualities that we take to be central or essential to who we are) is one that we sometimes share; in particular, we share it with our intimates.<sup>10</sup> And—importantly for my argument here—it is within certain loving intimate relationships that one can sometimes experience the kind of attention that can discover, reveal, and partly constitute one's identity and personality.

This attention, when provided by a loved one, will ideally partly crystallize and constitute our identities in ways that are consistent with our flourishing, more generally. As Amélie Rorty writes (pseudonymously, as Leila Tov-Ruach) of this relationship between love and attention,

When the lover's attentions are active in forming and crystallizing the beloved's personality, the lover is also careful to attend to the real structure of that personality, not foisting or projecting an identity that, by becoming constitutive, will so conflict with the rest of the beloved's character that the person cannot flourish. (Tov-Ruach 1980, 468)

Because of the roles that our loved ones play in discovering, forming, and maintaining our identities, we can come to experience their perspectives on who we are as possessing the epistemic, definitional, and evaluative credibility I described earlier. Most straightforwardly, we may acknowledge that our loved ones know us best. They tend to be the people we experience as sometimes having definitional authority over ourselves—an authority that is typically reserved for us alone. They attend to us in ways that can both reveal and crystallize our identities, leaving us with the immensely satisfying feeling of being “truly oneself” with and in the eyes of a loved

---

<sup>10</sup> For related discussion, see Elizabeth Spelman (1978). Spelman argues that there is a “maximal” sense of treating a person “as a person” that involves treating them in a way that takes into consideration their conception of themselves, as a necessary (but not indefeasible) constituent of their identity. She concludes that we should only expect a more minimal sense of treating a person “as a person” from non-intimates, given that our self-conceptions are aspects of ourselves that render us vulnerable, and which we are thereby inclined to keep private. I am suggesting that our intimates will be the ones most credibly suited to both take into consideration what we think about ourselves, as well as when and how we may be mistaken about this. Their credibility, along with their constitutive powers over our identities, is partly why we may become motivated to mask ourselves *especially* from our loved ones.

one. And because we may not only love them but also long for their reciprocated love, we also care what they think about us, and how they view us; we can, along with having our identities partly constituted by how they see us, feel diminished and injured by their perceptions of us in ways that the perceptions and attitudes of others may not matter to us at all, let alone partly constitute who we are. And importantly, we may not want our loved ones to view us in solely positive ways, or with good will or respect. This is because, more fundamentally, we may want them to view us as beings who are loveable by them – “mattering” in just this relatively narrow sense.

Return now to Morris’s reflections on the pain of wronging a loved one, and thereby losing or threatening one’s union, or relationship with her. As Rorty writes, within such a relationship it may become the case that “the person regards certain traits as centrally defining his personality and believes that he could not retain those traits outside of the particular attentional relation” (Tov-Ruach 1980, 468). Thus, this pain is not mere pain, but may be a response to the perceived loss of wholeness or completeness of one’s identity. And it may not be simply *painful*, but involve the disorientation, anxiety, and disassociation that results from one’s losing a sense of who one is. That we are vulnerable to certain others in this way is why acknowledging the perceptions and interpretations that they may have of us can come with an anxiety that may bear similarities to the anxiety felt in acknowledging our own biological deaths: we may not be merely hurt, disappointed, or angry when we are seen in certain ways, or when not attended to at all. We may feel as though we are losing a centrally defining part of, and so sometimes entirely, ourselves.

#### IV. From Love to Remorse, Remorse to Self-Hatred

Given this particular kind of vulnerability to our loved ones, we can now return to the question of how our love of others may render us vulnerable to self-hatred in the ways I’ve described, and which I’ve suggested aren’t best captured by thinking of the self-hating person as suffering from a guilty conscience. First, consider that love renders us susceptible to not only guilt, but remorse, which we become disposed to feel simply given our love of another, and given the possibility of being responsible for the damage or loss of the relationship that one shares with her. Unlike guilt, remorse need not be prompted by thoughts that one has violated a duty or obligation. Nor does it necessitate that one perceives oneself as morally responsible for

what has happened. But importantly, as with guilt, remorse will presuppose some form of responsibility, and so, will implicate oneself, and thereby give rise to, and sustain, hostility redirected onto oneself as the party responsible for this damage, or loss. The different sources of guilt and remorse account for differences in their phenomenology. On a Kantian conception of guilt, one feels both self-directed anger and anxiety, as a species of fear, at having violated an authoritative moral law. In feeling the worthy object of the anger of another, one may feel the pain of having taken something one has no right to, or of occupying a status that one does not deserve – one that that it would be justifiable for another to knock one down from. One may see in oneself a kind of self-conceit that calls for humbling, perhaps through the pain of deserved hard treatment.

Remorse, in contrast, is best understood as a form of sorrow, rather than as a form of fear. In remorse, one's attention may be focused more directly on the loved object that has been lost: in the cases relevant here, the beloved, her love, or one's relationship with her. If the union with her has been lost in a way where she has also suffered, one may keep returning to images of her, and her pain: as John Deigh (1996) suggests, a paradigmatic example of someone suffering from remorse is Vronsky, after the suicide of Anna Karenina. We see a man tormented not by the thought that he has violated a moral duty or obligation, but rather, by images of the person he loved, and whose suffering and ultimate death he perceives himself responsible, though not morally responsible, for. And given the role of our loved ones in maintaining our identities, one is again susceptible not just to the pains of separation (as well as the other forms of suffering mentioned above), but to the disorientation and disassociation of no longer *being oneself*.

In these respects, remorse is similar to another form of sorrow: grief. But unlike grief for a loved object that has been merely lost, remorse is felt when one perceives oneself as responsible for this loss. In remorse, the basic hostile responses that one would feel toward anyone who threatens or brings about the destruction of an object of one's love, one's relationship with the loved one, or the possibility of her love, can be redirected onto oneself as the responsible party – again, understanding responsibility here more broadly than “moral” responsibility. For example, consider the hostility and aggression that one might feel toward an actual other in jealousy. One perceives and feels that a rival has come to threaten an existing relationship of loving, mutual attention with a loved one, and feels hostile in response to the rival, even if she has not done anything to encourage the attention of one's beloved other than

just being, for example, her lovely and charming self. One may nonetheless view the rival as responsible for threatening that relationship, even when one holds no (insane) moral expectation that she not be as lovely or charming as she is.

Consider now when it is *oneself* who is viewed as the responsible party. And consider too, when the explanation why one can no longer sustain a union with the beloved, and her attention, has nothing to do with what one has done, but with what one is *like*. Whatever qualities one perceives as explaining why the beloved's attention has been lost or threatened can become the basis of one's hostility toward oneself: nonmoral aspects of one's character, one's physicality or style, background, race, or class-membership may become implicated in the explanation of the loss of this union, and so become the focus of one's hostility toward oneself. This is one reason why self-hatred can find its sources in such idiosyncratic aspects of the self, and why one can experience it as resulting from concerns that are only indirectly one's own. I may come to hate my lack of physical grace only once I imagine that it is what explains why I can no longer capture my beloved's loving attention—just as I might hate a rival's beauty for being what I imagine does capture it, instead. And this hostility is liable to intensify into hatred precisely in cases where what is threatened or lost is not only the relationship one has with a loved one, but also, one's own identity insofar as that relationship was identity-conferring. Given that the threat posed by another, or oneself, is experienced as a threat to who one is, or as responsible for a loss of one's sense of self, one's hostile feelings may intensify to the point where they may, as they do with hatred of others, culminate in a desire to destroy or eliminate the hated object.

Characteristics of remorse, when it develops into self-hatred, can also explain the self-hating person's willingness to accept and absorb hostility—whether that is her own hostility, or the hostility of the loved one, and so why self-hatred is liable to self-perpetuation. First, one's own suffering can take on a particular symbolic meaning in the context of the dissolution of a loving union, when one still has hope that the union might be restored. As Morris writes,

The satisfaction that one obtains in self-inflicted or accepted pain here comes from the very conduct as painful, for it is this that evidences... how much it means for one to restore... Indeed, part of what it means to love another, as well as oneself, is not only pain felt when the loved object is hurt, but pain one is prepared to face for the loved one. Therefore, in a genuine restorative response there may be a satisfaction derived from

restoring, a satisfaction derived from giving something to one for whom one cares, and a satisfaction derived from experiencing pain, for this makes apparent how deeply hurt one has been by the damage and how deeply committed one is to the relationship. (Morris 1971, 431)

While the pain that one seeks in guilt may symbolize how much one cares for one's *moral* relations to be restored with others, one's willingness to pay back one's debts, or an acceptance of humility, Morris here distinguishes from these forms of pain the pain that one is willing to undergo as proof of one's *love*.<sup>11</sup>

Second, another explanation becomes available in cases where one has lost hope that love will be returned or the relationship will be restored. In remorse felt because I have lost or can no longer maintain the beloved's loving attention, I may be motivated to retain just her attention, or what it has left in its wake. I may be willing to accept my (perhaps entirely) imagined (and perhaps exaggerated) perception that she has of me. At least, then, I will be able to retain some existence in her eyes, and thereby retain whatever aspect of my identity depended upon her, and her attention. Thus, *pace* Strawson, what we want from others cannot be limited to, or centered around, a desire that they have good will for us; *pace* Murphy, one may not solely, or most fundamentally, care for the "good opinion" or respect of another. When it comes to one's loved ones, what one may sometimes want is simply their attention, given how it may anchor and stabilize one's identity—even if that identity is one that they cannot also admire, respect, or love.

But importantly, self-hatred need not arise in the context of a failed relation to an actual other. Just as we don't need to be viewed by an actual other in a compromising way to feel shame, nor have our moral transgressions witnessed by an actual other in order to feel guilt, so we don't need to compromise an actual union with an actual other in order to feel self-hatred. Thus, we must return to questions about the various ways in which we relate to the internalized other. And in doing so, we can also begin to see—as we see with the self-hatred of Rufus Scott—how this other may come to drown out the voices of actual loved others who may actually love us.

---

<sup>11</sup> Another explanation, consistent with those that I offer is that a person, in fearing the perceived, likely hostility of an actual other, or an internalized other, will become motivated to *preemptively* experience that hostility, perhaps in the hopes of controlling how and when that hostility will be meted out. As Sandra Lee Bartky (1990, 89) observes (about certain rituals of self-induced shame), "An ordeal is often easier to endure if we can choose its time and place." My thanks for Francey Russell for discussion of this point.

To begin, consider Williams's remarks on the role of the internalized other in the experience of shame. He argues that we should not understand this other as reducible to an actual, identifiable other, nor are they identical to just my own voice and perspective. According to Williams, she embodies and focuses a real set of values and concerns that I am vulnerable to. I am vulnerable to her perspective of me given that I also respect her and hope for her respect. This provides us with the resources, he argues, to show how genuinely ethical motivation can emerge from sources that may initially seem too immaturely heteronomous to be ethical in nature, driving us, mistakenly, toward the Kantian picture of the autonomous moral agent instead. As he puts this point,

It is a mistake . . . to suppose that there are only two options, that the other in ethical thought must be an identifiable individual or a representative of the neighbours, on the one hand, or else be nothing at all except an echo chamber for my solitary moral voice... The internalized other is indeed abstracted and generalized and idealized, but he is potentially somebody rather than nobody, and somebody other than me. He can provide the focus of real social expectations, of how I shall live if I act in one way rather than another, of how my actions and reactions will alter my relations in the world about me. (Williams 1993, 84)

Consider, next, that it is not (or not just) that I respect this internalized other and hope only for her reciprocated respect, as Williams emphasizes; nor do I aspire (or just aspire) to become like her, as Velleman emphasizes. Another way in which I may relate to her is that she is an internalized figure of the kind of love that I have described, and whose reciprocated love I seek.<sup>12</sup> What she provides is not solely a focus of real social expectations and concerns, understood as a set of ethical or moral values. Rather, she also provides a focus of real expectations and concerns that determine whether or not a person is lovable by a hypothetical someone (an idealized,

---

<sup>12</sup> Marking the significance of this distinction between a hope for respect and a hope for love, Wesley Yang remarks of being Asian in the United States that "it was always the most salient of all facts, the one most readily on display, the thing that was unspeakable precisely because it need never be spoken: that as the bearer of an Asian face in America, you paid some incremental penalty, never absolute, but always omnipresent, that meant that you were by default unlovable and unloved; that you were presumptively a nobody, a mute and servile figure, distinguishable above all by your total incapacity to threaten anyone; *that you were many laudable things that the world might respect and reward*, but that you were fundamentally powerless to affect anyone in a way that would make you either loved or feared" (Yang 2018, 9–10, my emphasis).

generalized someone) whose love I would seek. She is not, as Williams puts it, an identifiable actual other, but nor is she just me. My standards of the lovability of others will inform who this idealized and generalized person is, but it is her standards, or what I imagine them to be, that I am vulnerable to—not because I have in any way *endorsed* them, but simply because I hope for her love. But just as with ethical standards, both my own standards and the standards that the internalized other embodies and focuses will largely be informed by the culture and society that I am raised in, and I need not grant these standards any kind of authority in order to be emotionally vulnerable to them.

Again, one may want from this figure both identity-conferring attention and love, and in happy cases, these do not greatly diverge from one another. But now imagine a less happy case. I realize that I am some way that compromises not just the internalized other's respect or admiration of me, but her love of me. I will feel the particular wound, the particular kind of unworthiness, hopelessness, sorrow and disorientation that one feels in not just being unloved by a particular and actual loved person, but as unlovable by *anyone* whose love I would care or need to have reciprocated. And my self-directed hostility will fixate on whatever it is I believe renders me unlovable to her: the idiosyncrasies of these qualities will depend on whatever standards of lovability have been embodied and focused by my internalized other.

To illustrate these points, consider how Pecola Breedlove in Toni Morrison's *The Bluest Eye* comes to hate herself for her darkness, which in the eyes of actual others marks her as "ugly." Given that the existing social standards of both the identity and lovability of girls and women heavily emphasize their physical appearance, she experiences her "ugliness" as both central to her identity, and as explaining why she is unlovable. She fixates on what she believes will make her beautiful, and thereby lovable, by those whose love she seeks: "the blue eyes of a white girl." And this is not just the love of actual others—she also sees herself as unlovable in the eyes of an idealized, generalized, and abstracted other, who, given that she is a child, is represented by her as an imaginary friend whose love and attention she longs for (Morrison, 203–4):

But suppose my eyes aren't blue enough?

*Blue enough for what?*

Blue enough for . . . I don't know. Blue enough for something. Blue enough . . . for you!

*I'm not going to play with you anymore.*

Oh. Don't leave me.

*Yes. I am.*

Why? Are you mad at me?

*Yes.*

Because my eyes aren't blue enough? Because I don't have the bluest eyes?

*No. Because you're acting silly.*

Don't go. Don't leave me. Will you come back if I get them?

*Get what?*

The bluest eyes. Will you come back then?

In my hopes to nonetheless achieve some sense that I am lovable by this internalized other, I may focus my energies on attempting to become so—sometimes with self-defeating and devastating results given that these standards will encompass much that a person simply cannot change about herself, because they are often conceptually inconsistent or otherwise practically impossible to meet, or because they conflict with what would make for a good human life.<sup>13</sup> One will be drawing from, after all, the imperfect materials of the social world that one is in, including the content of the fantasies of that world.

But we may also imagine someone who has lost hope. I simply cannot become the kind of person who—in the eyes of my internalized other—is loveable. Perhaps all I can do is inspire her contempt, or disgust. I may nonetheless strive for her identity-conferring attention, given that it holds the promise of maintaining a sense of my identity both in relation to her and in her eyes. Given this need, I may become willing to bear and absorb the unlovable self-conception that I imagine she has of me: again, at least someone contemptible, or disgusting to her, rather than no one, at least in relation to her, at all. My self-directed hostility will fixate on whatever it is I believe renders me unlovable: aspects of my identity that I imagine have cost me her love, but which may hold the promise, at least, of her attention.

---

<sup>13</sup> It may be a common experience that there will be no actual consistent standard here, or the standards will directly conflict with one another, or be impossible to achieve for other reasons. This will compound the difficulty one has in achieving a sense that one is *wholly* lovable, as the internalized figure will love only in irreconcilable and impossible ways. Pecola resolves this tension only through a delusional break from reality, in which she comes to believe that her eyes are in fact blue.

Of course, given that this identity is also one that one also hates—it is what one believes has rendered one unlovable, after all—this particular manifestation of self-hatred is liable to become unbearable: adherence to one’s identity will not entail an adherence to life. Return, finally, to the self-hatred of Rufus Scott. Among other aspects of himself, Rufus’s Blackness has become a source of his self-hatred. And it was his love for Leona which introduced in him a new kind of vulnerability that he had not, until then, experienced. In imagining how he is seen, now, as the Black lover of a white woman (abstracted and idealized), his identity as a Black man gains a new kind of evaluative significance for him: it has rendered him, in her eyes, and thereby in his eyes, unlovable. Spelled out with the resources that have been brought to bear, we can now see that his self-hatred persists partly because he cannot escape the eyes of this imagined, idealized and abstracted, internalized other whose love he longs for, but who he imagines cannot really love him, given his Blackness.<sup>14</sup> The gestures and expressions of the actual Leona who insists that she does love him, whatever they are intended to be, are now interpreted by Rufus through a lens: one partly constructed by his imagined identity in the eyes of this internalized other. It is not love that he then experiences her as expressing, but instead, a mix of condescension and sexual fetish:

“She loves the colored folks so much . . . sometimes I just can’t stand it. You know all that chick knows about me? The only thing she knows?” He put his hand on his sex, brutally, as though he would tear it out, and seemed pleased to see Vivaldo wince.<sup>15</sup>

Rufus is trapped by this self-conception, by his need to be both loved and seen by this internalized other in particular, and by his conviction that the person he is being seen as from that perspective

---

<sup>14</sup> Of the general need for recognition, George Yancy writes,

While I recognize the historical power of the white gaze, a perspective that carries the weight of white racist history and everyday encounters of spoken and unspoken anti-Black racism, I do not seek white recognition, that is, the [racist] white woman’s recognition. Though I would prefer that she does not see me through the distorting Black imago, I am not dependent upon her recognition. For me to seek white recognition as a stimulus to a healthy sense of self-understanding is a form of pathology. (Yancy 2008, 847–8)

The possibilities, routes, and repercussions of weaning oneself off certain forms of recognition are beyond the scope of this paper, though a clear and obvious benefit would be to escape from the self-hatred it, or particular forms of it (such as the recognition embedded in interpersonal love), can feed.

<sup>15</sup> (Baldwin 1992, 114, original ellipsis).

cannot also be one that is loved by her. Neither, however, is he able to continue living as the unlovable object of his own hatred. Tortured, Rufus walks onto the George Washington Bridge:

He stood at the center of the bridge and it was freezing cold. He raised his eyes to heaven. He thought, You bastard, you motherfucking bastard. *Ain't I your baby, too?* He began to cry. Something in Rufus which could not break shook him like a rag doll and splashed salt water all over his face and filled his throat and his nostrils with anguish. He knew the pain would never stop. He could never go down into the city again. He dropped his head as though someone had struck him and looked down at the water. It was cold and the water would be cold.

He was black and the water was black.<sup>16</sup>

---

<sup>16</sup> (Baldwin 1992, 144). My thanks to Kate Nolfi, George Sher, Francey Russell, Steven Klein, Ulrika Carlsson, Jonathan Gingerich, and Ram Neta. I am especially grateful to Robert Smithson, Samuel Reis-Dennis, and Oded Na'aman for discussion and comments on earlier versions of this chapter. Special thanks are owed to Noell Birondo, for his encouragement, comments, and editorial acuity. I am also grateful for discussions at The University of British Columbia, The Institute of Philosophy of the University of London, Boston University, Claremont McKenna College, as well as to participants and discussants at the 2021 Pacific Division Meeting of the American Philosophical Association, especially to Melissa Yates, for her delivered comments.