



Undergraduate Honors Theses

2024-03-15

**ON SENSIBLE LIMITS OF LOVE: IMPLICATIONS OF MORAL
FAILURE IN LOVING RELATIONSHIPS**

Macy West

Follow this and additional works at: https://scholarsarchive.byu.edu/studentpub_uht

BYU ScholarsArchive Citation

West, Macy, "ON SENSIBLE LIMITS OF LOVE: IMPLICATIONS OF MORAL FAILURE IN LOVING RELATIONSHIPS" (2024). *Undergraduate Honors Theses*. 361.
https://scholarsarchive.byu.edu/studentpub_uht/361

This Honors Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.

Honors Thesis

ON SENSIBLE LIMITS OF LOVE:
IMPLICATIONS OF MORAL FAILURE IN LOVING RELATIONSHIPS

by
Macy West

Submitted to Brigham Young University in partial fulfillment
of graduation requirements for University Honors

Philosophy Department
Brigham Young University
April 2024

Advisor: Justin White
Honors Coordinator: Ryan Christensen

ABSTRACT

ON SENSIBLE LIMITS OF LOVE: IMPLICATIONS OF MORAL FAILURE IN LOVING RELATIONSHIPS

Macy West

Philosophy Department

Bachelor of Arts

In this thesis, I aim to establish the groundwork for an account of love that is sensitive to serious moral failure, both descriptively and normatively. I attempt to articulate not only how love is compromised by abuse, but also by serious moral failure that is not directed at the relationship participant. I first refute two popular accounts that argue for the compatibility of moral failure with love, Vida Yao's permissive love and Eric Matthes's dismissive love, on the grounds that they fail to adequately respond to more serious moral flaws. I then propose my own conception of practical identity as it pertains to moral flaws and draw an essential distinction between benign and malicious moral flaws. I posit that malicious moral flaws are, in almost all cases, incompatible with a genuinely loving relationship due to the estrangement that they foster. I argue that under a Kolodnian view of love, we can meaningfully rescind the love we've extended to a relationship. Finally, I claim that the vulnerability and morally fraught dispositions that result from engaging in a relationship with a maliciously flawed person imbue us with a responsibility to disengage from that relationship.

ACKNOWLEDGMENTS

I would like to thank my parents, Beca and Andy, and my fiancé, Adam, for showing me the most exceptional examples of genuine love that I could ask for. I also thank Ryan Christensen, my Honors Coordinator, for serving as a liaison between the Philosophy Department and the Honors Program despite geographic difficulties. My faculty reader, Derek Haderlie, deserves special mention for reassuring and encouraging me not only throughout this thesis project, but my entire undergraduate education. Finally, I express my deepest gratitude to my faculty advisor, Justin White, for convincing me to major in philosophy four years ago, inspiring the topic of this thesis, spending countless hours providing detailed feedback, and everything in between. I could not have written this thesis without him.

TABLE OF CONTENTS

Title	i
Abstract	iii
Acknowledgments	v
Table of Contents	vii
I. Introduction	1
II. On Permissive Love	5
III. On Dismissive Love	8
IV. Defining Moral Flaws	14
V. Theoretical Implications: Estrangement and Attachment	21
VI. Practical Implications: Vulnerability to a Relationship	30
VII. Conclusion	43
Works Cited	45

I. INTRODUCTION

*“When you love someone, you love the whole person, just as he or she is,
and not as you would like them to be.”*

— *Leo Tolstoy, Anna Karenina*

The sentiment expressed above is certainly pervasive in more than just classic literature. Indeed, we want to be loved for our whole selves; we want to be accepted and appreciated not only for our exemplary characteristics, but also in light of the objectionable, and even disgraceful ones. Given this, we might think that we ought to strive to love others in the way we hope to be loved. But could it be problematic to love people with morally reprehensible qualities? Vida Yao (2020) would answer negatively, arguing that we should extend grace to those we love, which requires affection towards the beloved’s failures since they are a part of their human nature. There is something intuitively appealing about this account that speaks to a desire many of us have—that is, to be loved just as we are. But there’s also something differently appealing about the notion of being able to “look away” from our loved one’s flaws, and focus on their best qualities. In this vein, Erich Matthes (2016) argues that we should love a person *in spite of* their flaws, but not *for* them. Matthes takes issue with the notion of overly permissive love, and instead suggests that we can and should exclude certain questionable traits from our motivational reasons for loving a person.

Yao and Matthes are not just interested in particular instances of moral failure (or immoral action); they are more concerned with addressing qualities that constitute *moral flaws*, or qualities which manifest through immoral behavior. To qualify as a moral flaw, the behavior resulting from the flaw must be sufficiently pervasive to the extent that it is a part of the person’s identity. Moral flaws, which I will further elaborate on in what follows, are distinct from any one

occurrence of moral failure. Yao's and Matthes's perspectives illustrate two possible approaches we might apply to deal with moral flaws we identify in our loved ones: loving their flaws as constitutive of their endearing humanity (a more permissive approach), or ceasing to consider those flaws as a part of their identity when loving them (in other words, dismissing the flaw from our reasons for love). But are either of these responses sufficient in cases of severe moral flaws? Should a woman who discovers her lover poisoned innocent children for financial gain continue to love him for, or even in spite of, that flaw? What about a woman whose brother physically abuses her? These are just a couple of the examples I will examine in this thesis to explicate the problematic nature of loving maliciously flawed people. I will argue that both Yao's permissive and Matthes's dismissive approaches fail to provide an account of love that can adequately respond to certain types of moral flaws.

I believe that the complexity of the intersection of love and moral failure has been given short shrift in the philosophical literature of love.¹ To be clear, the notion that victims of abuse have reason to leave a relationship or cease to love their abuser is not foreign in the philosophy of love. But I want to articulate precisely *why* this is the case, and what aspects of love are compromised as a result of abuse. I also want to make the further claim that not only do we have reason to disengage from relationships in which our beloved acts maliciously towards us, but that we also have a reason to disengage from relationships in which our beloved's malicious behavior is not directly harmful to us. Moral flaws directed at a relationship participant (such as abuse) are

¹ This is not to say that discussion of love in the context of abuse and/or moral failure is nonexistent. For instance, "Love, Incorporated" by Adrienne Martin (2015) addresses the notion of ambivalent love and questions how an abusive person can simultaneously abuse and love their partner. In "The Illusion of Love: Why the Battered Woman Returns to Her Abuser," David Celani (1994) examines why and how cycles of abuse are perpetuated and the role of personality disorders in these cycles.

a central focus of this project, but I am also interested in addressing how malicious behavior outside the realm of a relationship can and should affect the relationship.

I aim to charitably represent Yao's and Matthes's theories, as they are compelling and useful in many cases. In sections II and III of this thesis, I hope to illuminate why popular suggestions for addressing moral flaws in loving relationships do not provide a satisfactory response for severe moral flaws.² Before we can understand what is required for a robust account of moral flaws in loving relationships, we must first define moral flaws and distinguish relevant differences within the umbrella term. In section IV, I will suggest a simple two-way distinction: benign and malicious moral flaws. I will also posit that moral flaws "in" loving relationships may refer to either moral flaws directed at the relevant relationship participant or moral flaws directed at others, as they are both vital to evaluating the status of a relationship.

In section V, I will argue that the known presence of malicious moral flaws in a beloved invariably compromises the status of a relationship as loving. This is because, once discovered, the presence of a malicious moral flaw in the beloved necessarily leads to estrangement and oftentimes an inability to see the beloved as they truly are, both of which are contrary to popular requirements for love. I will claim that the pseudo-loving status of a relationship with a maliciously flawed person is, in many cases, better described as attachment.³ Adopting Monique Wonderly's definition of attachment, we can see that attachment involves a desire to engage with

² I am exclusively interested in loving relationships, in which loving attitudes are at least somewhat reciprocated and the two moral agents have a relationship that persists over time. Unrequited love and "love" that people may feel towards celebrities or leadership figures that they do not know personally are outside of my project's scope.

³ I will use the phrases "maliciously flawed person" and "malicious person" interchangeably for the purposes of this project.

the beloved, and a sense of loss when they are absent, which are characteristic of close relationships—even ones that do not meet important criteria for genuine love.

However, even if the theoretical status of genuine love is compromised by abuse and immorality, these circumstances do not magically extinguish our loving feelings towards the beloved. We sometimes find ourselves “loving” someone despite our deepest wishes to cease doing so. So what does it mean for the presence of a malicious moral flaw to invariably compromise love if we are unable to rid ourselves of the loving attitude we have towards the beloved? I will respond to this question in section VI by outlining a relationship-based account of love to address the practical implications of moral flaws. Using Niko Kolodny’s (2003) theory of love as valuing a relationship, I will attempt to show that disengaging from a relationship is effectively withdrawing love even if there isn’t an immediate change in loving attitude. Beyond the mere possibility of disengagement, I want to argue that we *ought* to disengage in a loving relationship once malicious moral flaws become apparent on the grounds that continued vulnerability to a maliciously flawed person can endanger victims of abuse and lead to harboring damaging dispositions.

As I alluded to above, one of my primary motivations for investigating moral failure in the context of love stems from a desire to give abusive relationships the philosophical attention they deserve. The philosophy of love is largely concerned with discerning the nature of love, and as such, tends to err on the side of ideal love—we are interested in metrics and criteria that ideally constitute love. But love is rarely ideal, and unfortunately, abuse and immorality in loving relationships is common enough that it’s important for us to have an account of love that is sensitive to the way love can be undermined by moral flaws. Considering how frequently abuse and love do come into conflict, if we really seek to explain the phenomenon of love, we

ought to consider how the nature and status of love changes in the face of malice, and what we are to do if that does happen, practically speaking. I hope to provide the groundwork for such an account in this thesis. I believe it is possible (and important) to craft an account of love that provides victims of abuse, and more generally, agents who love malicious persons with some control over the status of their love.

II. ON PERMISSIVE LOVE

In her paper “Grace and Alienation,” Vida Yao emphasizes the importance of grace and attentiveness for a robust account of love. She believes that in order to *truly* and ideally love someone, we must overcome the alienation caused by shame (5). Yao defines shame as the painful experience of being perceived in the wrong way, either by another person whose perception matters to the agent, or by an internalized “other” (4). Shame fosters alienation because when a person is ashamed, their entire sense of self is diminished, and they often wish to hide or disappear. This also causes them to estrange themselves from their loved ones, which is antithetical to love. It is only when this estrangement is repaired that genuine love can be fostered (5). To aid our beloved in overcoming shame, we must extend affection to all the characteristics of the beloved, especially those not considered “good” (15).

Inspired by Iris Murdoch’s work, Yao begins by endorsing two important criteria for love: first, love first fosters connection and communion between the involved parties rather than estrangement; and second, love is attentive, with the ambition of this attention being to see the person as they truly are (Yao 2). Yao is especially concerned with the second criterion and how being attentive to the beloved strengthens and constitutes a loving relationship. Contrary to popular wisdom that we should be less attentive to our loved one’s flaws (e.g., “love is blind”), Yao advocates for a conception of love in which we specifically attend to those flaws and work

to love them *for* those flaws rather than in spite of them. In her words, we should cultivate an “affectionate love felt in response to perceived qualities of human nature” and “become endeared” to the beloved “upon noticing [their] possession of such qualities, in virtue of recognizing [their] possession of those qualities” (15). In the context of this passage, she is referring specifically to qualities commonly labeled as “bad” or “evil” and nonetheless advises that we become endeared to the beloved *by virtue* of their very worst qualities. When we correctly attend to the beloved, we not only strive to be aware of all qualities of the human nature of the beloved, but also to perceive them in an endearing light. Importantly, we do not base our love on a wish for the beloved to change, or the hope that they can improve (11). We attend to the flaws *themselves* as we learn to be affectionate towards them (16). Without attentiveness, we fail to meet Yao’s aforementioned criteria of seeing the beloved as they truly are. But is attentiveness alone enough to fulfill both criteria? Perhaps the more attentive we are to our loved one’s flaws, the more repelled by them we will be, thereby estranging ourselves from them and failing to meet the first criteria. Indeed, it may seem unintuitive, or even shocking, to expect that we could actually feel affection or love towards morally reprehensible qualities in our loved ones.

In order to make attentive love possible in this way, Yao proposes that we must supplement our attention with grace. Grace, according to Yao, is “non-proportional to the excellence of its objects” and, as a result, can allow us to exercise a love that is responsive to all qualities of the beloved, including the “bad” ones (3). Only with attentiveness *and* grace can we foster intimacy and communion with the beloved without shying away from their flaws. To illustrate this, Yao uses a literary example from Marilynne Robinson’s novel *Home*, outlining a conversation between Glory and her brother Jack (Yao 10). In it, Glory tells Jack that she likes

his soul just the way it is. Jack finds this impossible, claiming that she must be mistaking him for someone else. He begins to list all the ways in which he is flawed and therefore unlovable: he's a drunk, a thief, takes advantage of vulnerable women, is incredibly vain, malicious, etc. As he articulates his list, Glory nods in agreement, expressing that she is already acutely aware of all of these flaws. In doing so, she confirms to Jack that she is seeing him as he truly is, and she likes his soul, including all of his worst flaws, without expectation or desire that it change or improve (Yao 10). We see that Glory has cultivated affection for Jack's mischievous qualities, because they are what make him human to her. With this example, Yao effectively illustrates how gracious, attentive love is needed to break down barriers of shame in loving relationships. It is not merely that Glory is affectionate to Jack without regard for his flaws—she feels affection towards Jack *in light of* his flaws.

Yao thinks that this sort of affection is possible *only* when we see flaws as endearing traits of human nature. She uses another example of a woman named Jane who, in an act of hatred and jealousy, tears up a photo of her ex-friend Joan and gouges out the eyes (15). Yao sees this act as endearing, and even touching, despite its unprovoked, violent nature. She believes the recognition of flawed, and even unreasonable behavior like Jane's eye-gouging can be delightful in its relatability and prompt affection for the humanness of our loved ones. After all, flaws make humans intelligible to one another. Without moral failure, our loved ones would be eerily alien to us, so it makes sense that, with grace in mind, we could be endeared towards reprehensible qualities of the beloved.

There is certainly valuable insight to be gained from a close reading and application of Yao's theory of love. We can, and should, strive to allow our loved one's flaws to endear them to us if we are to truly love them as they are. Yao does an excellent job of motivating the

importance of rooting out shame and introducing a permissive, accepting form of love. But her feel-good, permissive approach belies darker, more concerning implications. What if we were to change the circumstances of the stories Yao uses? What if Jack was a rapist, or a murderer? What if Jane's eye gouging was not confined to paper photos, but directed at real people? It seems unreasonable to suggest that in such severe cases of moral failure, we should be affectionate towards these behaviors or find them endearing. And yet Yao places no explicit limit on the extent to which she advocates for grace and affection. She does not specifically address how or if more serious moral flaws affect her conception of grace and attentive love, so it is unclear what her stance on continuing to extend grace towards truly egregious qualities would be.

Yao's account seems to imply that we should be permissive of and extend love towards even truly egregious moral flaws in our loved ones. It is possible that this implication could be remedied by a clarification of the extent that Yao's grace should be applied. But drawing a particular line in the sand for an approach that seems to rely on unlimited gracious permissivity would surely prove difficult. At best, her account is simply incomplete in that she fails to address necessary limits to grace; at worst, it is a dangerous account that provides fodder for justifying remaining in dangerous and abusive relationships.

III. ON DISMISSIVE LOVE

In the previous section, I attempted to show that the permissive love approach for moral failure in relationships does not adequately consider more severe moral failure. I will now consider another approach for responding to moral failure in loving relationships, which I will coin the "dismissive approach." In his essay "Love, in Spite Of," Erich Matthes argues that we should respond to moral failure by excluding, or dismissing the failure from our motivational

reasons for loving the person.⁴ He begins his argument by assuming a conception of love in which love is a *valuing attitude* (242). In other words, according to Matthes, loving someone is to have a range of positive attitudes towards them that are viewed as warranted in virtue of certain facts about a person, otherwise known as reasons for love (242). Matthes proposes two primary criteria for loving a person, which he calls the identity requirement and the well-being requirement. Although these criteria differ from Yao's, they are largely similar and are likewise inspired by Iris Murdoch's concern with attentiveness. Matthes's identity requirement posits that to actually love a person, we must be properly responsive to a person's identity by loving them for who they *truly* are (243). Matthes posits that we can determine who a person truly is by attending to their practical identity, or the description under which they value themselves (243).⁵ He further clarifies his definition of practical identity to mean who a person *is* "evaluatively, by [one's] own lights" and "how [one] conceives of [oneself] as belonging to a certain social category" (243). For instance, Jack explicitly evaluates himself as vain and morally fraught. This indicates that he sees these qualities as a part of his identity. Matthes's second criterion (the well-being requirement) requires that we possess and act upon a general concern for the beloved's actual well-being, to the best of our understanding (244).

But within this notion of love, Matthes points out a common conflict: what are we to do when loving a person for who they truly are involves loving characteristics that actually harm their well-being? Jack would be better off if he was not vain and malicious, so if Glory loves Jack for who he truly is, and cultivates affection towards his flaws, she loves him *for* qualities

⁴ A large portion of Matthes's argument in this essay pertains to loving people with physical and mental disabilities. I have excluded this section from my discussion as it is independent from his argument regarding moral failure and is outside of the scope of this project.

⁵ Matthes adopts Korsgaard's view of practical identity here.

that harm him. If Glory is to fulfill the well-being requirement for love, she should not feel affection for qualities that go against his well-being. And yet, if she is to fulfill the identity-requirement, she must find a way to love those harmful qualities. It seems these situations pit the two requirements for love against one another. It is important to note that under Matthes's account, all immoral actions are harmful to a person's well-being, even if they do not directly harm the person themselves. This is due to the fact that moral wrongdoing inhibits one from flourishing as a person, which in turn harms their well-being (246). Although the claim that all moral wrongdoing is damaging to one's well-being may seem radical, Matthes believes that it is actually a rather intuitive claim if we grant that behaving morally is important to flourishing in society and as an individual. As such, the presence of moral wrongdoing (either intentionally or inadvertently self-damaging), if sufficiently pervasive, indicates that moral flaws are present in a person's practical identity.

Matthes's proposed solution for responding to situations in which moral flaws are present in a person's practical identity is to love them in *spite* of the wrongdoing. To illustrate what this looks like, Matthes provides an analogy: imagine that a beautiful sculpture is damaged and vandalized with graffiti. We may still value the sculpture for its artistic worth but exclude the graffiti entirely from our reasons for valuing it as such (249). Likewise, we can still love a family member who exhibits racist attitudes but exclude that behavior entirely from our reasons for valuing them in a loving manner. Matthes contends that it seems wrong to suggest that we should love the family member for who they are and, as a result, love them *for* being racist (247).

Matthes grants that loving someone in spite of part of their practical identity necessarily compromises the identity requirement, since we intentionally cease to see the person as they truly are by quietly dismissing their flaws. He does not see this as a problem for his solution. In

fact, Matthes sees this as a desirable feature of his view. He claims that if we were to truly see our loved ones for who they are all of the time, love would be impossible:

If we couldn't love people in spite of some things, then we couldn't love people at all. I'm pretty sure no one would love me, at least. We might actually think that loving in spite of is love's natural state. It's what makes love possible. To insist otherwise would be a bizarrely puritanical and unrealistic picture of loving behavior, requiring the simpering endorsement of every obnoxious quirk our beloveds possess. (252)

Matthes's point here marks a near-direct refutation of the permissive approach, as he asserts that without excluding certain characteristics from our love, we simply couldn't be expected to love anyone. People are imperfect and love, as a result, is imperfect as well. With this in mind, Matthes believes that we can still genuinely love a person even when we love them in spite of certain qualities, and fail to perfectly fulfill the identity requirement.

Matthes's account does address some of the concerns I raised with Yao. He recognizes that it may be inappropriate to feel affection towards moral failure, and his theory allows us to essentially "dismiss" moral flaws while still being attentive to the bulk of a person's practical identity. This way, we can avoid the unsettling suggestion that we ought to cultivate affection towards worrisome flaws. But although Matthes account is less permissive than Yao's account, it still belies the same concerning implications. In what follows, I will argue that Matthes's account relies on an implausible version of practical identity, and in application, leads to a dangerous avoidance of malicious behavior.

As previously mentioned, Matthes specifies early in his essay that one's practical identity is made up of traits "without which you wouldn't be *you*, . . . evaluatively, by your own lights" (243). This definition is broadly Korsgaardian in nature; Korsgaard influentially defines practical

identity as a “description under which one values oneself” (101). Under Korsgaard’s view, we find ourselves with impulses and desires that we either reject or endorse. But to do so, we need some metric to assess these impulses. Practical identity provides us with the point of view by which we can assess desires that arise. We adhere to certain descriptions of ourselves when determining our values.⁶ While Korsgaard would not necessarily grant that we are the sole constitutors of our own identity, she does place significant emphasis on what we as selves decide is important to us. Matthes supports a stronger version of this idea, endorsing the claim that our identity is entirely up to us. He articulates practical identity, at least as far as it pertains to love, as a product of the way we view ourselves, and what we take to be important about the self we are (find page). There is certainly something compelling about this view. If love requires that our practical identity must be attended to, and we are the determinors of that identity, we have full disclosure in determining which of our qualities should be seen and attended to by our loved ones.

However, as appealing as the notion of self-creation might sound in this context, it is sorely inaccurate given the difficulty of accurate self-articulation. We are all likely guilty of such inaccuracy, at some time or another. Perhaps we like to imagine we are a little kinder than we truly are, or a little more innocent than we truly are. Or in the opposite direction: a little more pathetic or unsuccessful than we truly are. Depending on the person and the situation, we may over or underestimate the prevalence of certain qualities in our practical identities. To be fair to Matthes, a self-constituted version of practical identity does elegantly capture the sense in which our self-conceptions play a crucial role in our agency and the possibilities we open to ourselves.

⁶ For instance, I might describe myself as kind and athletic, and those descriptions will likely lead me to choose to value behaviors and/or activities that are charitable, or that promote physical activity.

But unfortunately, to accept a version of practical identity that is permissive of genuine errors is to accept an inaccurate picture of who we truly are. This move highlights some unsettling implications for Matthes's argument, seeing as it allows the beloved to write off moral flaws as not really theirs? so long as they do not understand them to be a part of their own practical identity. Thus, before we can address other issues with Matthes's argument, we must first push back on the implicit assumption of practical identity as wholly self-determined that undergirds his entire argument. If we were to accept Matthes's view of practical identity, we would be unable to focus on genuine moral flaws hidden from the first-personal identity through self-deception, leaving us with a significantly less robust picture of how moral flaws function in relationships.

I will now address my secondary issue with Matthes's argument. Let's imagine that Matthes alters his definition of practical identity to account for third-personal perspectives by granting that practical identity is not entirely self-constituted, thereby circumventing my initial objection to his essay. Would his account then be an effective way of responding to moral failure in loving relationships? Imagine that you are in a relationship with someone who acts violently towards you and others. Perhaps they occasionally lash out and hit you when they are angry. Or perhaps you remain physically unharmed, but are informed that your partner assaulted a coworker when angered at work. When a trusted friend confronts you about the situation, you respond with the following statement: "I love my partner in spite of their violent behavior. I just don't consider it when I think of reasons I love them. After all, they have many other lovable qualities to focus on." Your friend would likely find something deeply wrong with this statement, and as a more neutral third-personal perspective, they would be valid in doing so. The

friend's concern in this scenario aptly highlights a worrisome element of loving in spite of in some cases.

Excluding severe moral flaws from the motivational reasons for loving someone fails to address the root of the issue. When we become aware of serious moral failure on the part of a loved one, we ought not to merely avoid it or dismiss it. In fact, this course of action has the potential to be even more dangerous than Yao's attentive love, seeing as it may lead to perpetual concealment of the true depth of a flaw in the name of loving a person "in spite of" that flaw. If the solution for loving imperfect persons is to automatically exclude their worst qualities from evaluation, how are we to evaluate the severity and damaging nature of those qualities? Setting such an expectation opens the floodgates for avoidant behavior in unhealthy relationships.

IV. DEFINING MORAL FLAWS

Now that I have addressed the two most prominent views for responding to moral failure in loving relationships and highlighted what I've argued are fatal flaws in both accounts, I will provide my own suggestion for responding to more severe cases of moral flaws. But before I can do so, I must provide a descriptive account of moral flaws themselves. In this section, I will outline a revised conception of practical identity as it pertains to moral flaws. I will then propose a two-way distinction between benign moral failure and malicious moral failure.⁷

In the previous section, I rejected the notion of wholly self-constituted practical identity. Especially for the discussion at hand, we need a conception of practical identity that protects against self-deception and includes traits that may be hidden from the first-personal perspective (as a result of self-deception). More specifically, I argue that the resources that a partially *third-*

⁷ One might argue that if a behavior is a genuine moral flaw, it cannot be benign by definition. It is important to note that I am using benign as a technical term here, which I will define in my own terms, so one need not be concerned that I am claiming benign moral flaws are not harmful whatsoever.

personal view of practical identity affords us allow us to better capture the nuances of potentially abusive relationships.

So then, if practical identity is not just self-determined, who gets to determine it? I argue that practical identity (or the self that we truly are) should be determined and corroborated first-personally *and* third-personally.⁸ There are vital insights into the people we truly are held in our relationships with other rational agents. Whether it be from new acquaintances or close friends and family, the way we are perceived by those around us, excluding those who are likely to have a distorted perception (such as a grudge or quarrel that clouds one’s perspective), is a good indicator of the way our “true self” actually manifests in the world.

If we find that our self-perceived practical identity is significantly at odds with our third-personally perceived practical identity, this should give us reason to pause, and re-evaluate the possible presence of self-deception (whether intentional or unintentional) that is clouding us from being aware of the person we truly are. But who gets to be a part of this vague “third-personal” group of agents? To draw on a passage from Nietzsche’s *On the Genealogy of Morality*, it matters more that we consider the views of many eyes and different eyes than that we identify a set of particular eyes to give a voice to (3:12, 87). Unless one has a reason to discount some particular perspective, each third-personal perspective should be given genuine weight and consideration for the insight it can provide into the selves we truly are.

Clarifying my concept of practical identity may seem less relevant to the discussion of loving relationships, but given that my argument addresses *moral flaws*, not just any moral

⁸ I do want to note that even this version of practical identity is imperfect. Of course, there exists some “objectively” accurate identity that could, in rare cases, be concealed from both the first-personal and third-personal perspectives. But in the vast majority of cases, someone from a third-personal perspective would have an accurate perception of at least a part of a given person’s identity.

failure, a robust definition of practical identity is a prerequisite to further discussion. When I state that a person has a “moral flaw,” I mean that a certain immoral behavior is pervasive enough to qualify as a part of the person’s practical identity. It is, in some sense, an integral part of who they *truly* are, whether or not they would agree. If I verbally lash out at my sibling one time, this is certainly a moral failure, but it’s less clear that it would be a *moral flaw* that is a part of my identity. If I lash out at my siblings and other loved ones very consistently, this behavior is a likely candidate for a moral flaw, since it is a relevant part of who I am as a person in some deeper sense.

However, consistency isn’t necessarily required for a moral flaw to exist. If I murder an innocent person only once, the action is so severe that it, even in only one instance, clearly indicates something integral and vitally relevant about the person I am. This is because certain immoral actions are so extreme that only someone that inhabits a certain disposition or quality would ever see fit to commit that action. A “typical” moral agent would not suddenly find themselves resorting to murder; even in crimes of passion, there seems to be something important to understand about a person through these actions, even if they are outliers. The response I intend to advocate for does not quantify over instances of moral failure. I am concerned with behavior that is relevantly pervasive and/or important to a person’s identity. I state this now to clarify why first and third-personal perspectives are vital to a concept of practical identity that captures the full “story” for a person’s moral flaws in loving relationships, and, by extension, vital to this project.

Now that we have established both that moral flaws can be determined first-personally or third-personally and that they must be a relevant part of who a person is to qualify as a flaw rather than a mere failure, we must further investigate possible differentiations within moral

flaws. For the purposes of this paper, I will focus on one distinction: one that divides between benign and malicious moral flaws. Moral flaws range from the minorly immoral (benign) to the severely immoral (malicious), and it would seem theoretically misguided to make no distinction between the two ends of the spectrum. It is not quite the same to say that we can love a friend despite their flakiness or abrasive personality as it is to say that we can love a friend despite their history as a murderer. While it is surely difficult to draw a hard line on this spectrum, some coherent distinction can be made. The line between benign and malicious moral flaws may be vague, and how we draw this line may vary among different ethical theories, but this is not problematic for my project, as I am primarily interested in addressing moral flaws that err very clearly on the side of malice.

To construct a method for determining whether a moral flaw is malicious or benign, I will consider two main factors: (1) how the agent relates to the flaw and (2) the severity of harm caused by the flaw. These factors work together to determine the maliciousness of a flaw. There are a variety of ways an agent might react or relate to their flaw. They may be disgusted with themselves and resent the flaw, or alternatively be proud of the flaw. They could also have a perfect awareness of the flaw, or be entirely oblivious to the flaw and the accompanying behavior. The more ardently positive an agent's relationship with their flaw is, the more likely it is to be malicious. An agent would have an obviously positive relationship with their flaw if they (1) are aware that it is a flaw and relish in it nonetheless or (2) recognize the behavior associated with the flaw and believe it to be genuinely valuable or good. These two situations are admittedly quite different, but both display a positive relationship with the flaw. If an agent is entirely unaware of the moral flaw, it may be harder to categorize how they relate to the flaw. This does not mean the flaw is not malicious, it just means that we will need to turn to the latter factor of

severity of harm to determine this. The confusion that lack of self-awareness can introduce for categorizing moral behavior illustrates why my early revision to the concept of practical identity is so essential.

For instance, let's say I have a friend, Sue, who is incredibly flaky: she bails on plans last minute, sometimes without ever notifying me, and often fails to fulfill important promises she has made. Imagine that she is entirely unaware of this behavior—she's oblivious to the frequency or severity of her inconsistency. When the behavior is brought to her attention, she is deeply apologetic and vows to improve despite a general failure to do so. This flaw, although clearly pervasive and inconvenient, does not necessarily strike me as malicious because of the way that Sue relates to the failure: she was previously unaware of it, and once made aware, was genuinely apologetic and attempts to change the behavior, albeit with low success. But if we alter the scenario to posit that Sue is fully aware of her flakiness and of the adverse effect that it has on her loved ones, and knowing this, is proud of the control she wields over others by causing them harm, this seemingly benign behavior becomes malicious. A positive relationship with one's flaw is obviously a good indicator that malice may be present, but it does not guarantee that the flaw is malicious. If I have a positive relationship with a relatively minor moral flaw, it may not be severe enough to warrant being classified as a malicious flaw. This brings me to the next factor.

The second factor, severity of harm, is more concerned with the nature of the flaw itself rather than the agent's relation to the flaw. Although a desire to distance oneself from the parts of one's identity that are morally reprehensible is a good thing, it cannot always prevent the flaw

from being malicious.⁹ If the harm caused by a moral flaw is sufficiently severe, the flaw is malicious. By harm, I mean physical or mental damage/injury that is sustained over some period of time.

Importantly, the metric I am using to assess harm considers the harm that would occur if a concealed offense was revealed to the relevant relationship participant. The phrase “what they don’t know can’t hurt them” is often used to justify malicious flaws, such as sustained cheating on a romantic partner or other forms of relationship betrayal.¹⁰ Just because one’s partner is unaware of an affair does not mean that there is not great harm done to the partner. Thus, we should consider actual damage done rather than the perception (or lack of perception) of a harm.¹¹

Of course, as mentioned above, this harm can be physical or mental in nature.¹² Some level of mental distress or damage will come as a natural result of many benign moral flaws in a loving relationship, as humans are imperfect and likely engage in minor immoral behavior towards each other somewhat frequently. But significant, lasting, and/or irreparable mental damage is not a normal result of a loving relationship. Furthermore, there are few, if any, cases in which any physical harm should be inflicted on either participant in a loving relationship,

⁹ I understand that this use of malice breaks from the textbook definition, so I do want to note that I am using the term malicious as a technical term, which is defined solely by the terms set in this project.

¹⁰ Although I am staunchly against affairs, I do not believe every affair constitutes a malicious moral flaw that is present in the offender’s practical identity. For the sake of this quick example, I will assume this case is a malicious moral flaw.

¹¹ If the lover is unaware of the affair, and the beloved is in denial about the qualities they possess that led them to betray their partner, this could be one of those rare cases where neither the first-personal nor third-personal perspective accurately captures the practical identity of the beloved.

¹² Understandably, there are nuances within the mind/body distinction that the physical/mental distinction rests on and I am surely oversimplifying them. By “mental” damage, I simply mean any sort of damage that is not inflicted on a person’s physical body.

barring genuine accidents. Even an action as seemingly minor as inflicting small scratches on a loved one when angered constitutes a malicious moral flaw. Any level of pervasive and intentional physical harm introduces a dangerous power dynamic into a loving relationship that leads to lasting mental damage. When the possibility of physical harm, however slight, characterizes a loving relationship, trust and openness are replaced with fear and uncertainty, with the inflictor of the violence holding power over their relationship partner through fear.

Another important nuance to clarify is that the harm does not have to be inflicted on the relationship participant to qualify as a relevant malicious moral flaw. When a loved one inflicts serious harm on others, even if you remain unscathed, their moral flaw still bears on *your* relationship with them. If my friend is physically abusive to her spouse, but never to me, that malicious moral flaw is still deeply relevant to our loving relationship. To engage in a typical loving relationship with someone is, in part, to endorse and support them, their actions, and their vices. I will provide a more robust argument for the relevance of these sorts of flaws to relationships in what follows, but suffice it to say, one cannot maintain a relationship with a maliciously flawed person (even when the flaw is directed elsewhere) without either estranging themselves from their beloved, or becoming complicit in their wrongdoing. Thus, immoral actions, even when not directed at participants of a certain relationship, are relevant to the status of that relationship.

I have used the terms “abuse” and “abusive” semi-frequently up to this point. Now, furnished with a more careful conception of moral flaws, I will be more specific as to my use of these terms. Abusive relationships, I claim, are relationships characterized by malicious moral flaws directed at the relationship participant. By tying the popular notion of abuse with a more

specific type of behavior (a combination of the severity and agent's relationship to it), we can better explore the implications that malicious behavior has for love.

I have previously argued against both Yao and Matthes's theories of love on the grounds that they cannot account for how we ought (or ought not) to love in the face of malicious moral flaws. The issue with both of these accounts is that they provide inadequate responses for relationships that involve malicious moral flaws. Both of these accounts are likely defensible when applied to benign moral flaws, which occur in all relationships to some extent. But a more drastic approach is appropriate for malicious flaws and abusive behavior in relationships. Before I propose my preferred approach, I want to highlight how malicious moral flaws in the beloved impact one's ability to connect with the beloved. In the following section, I will argue that "loving" a person with malicious moral flaws necessarily fails to meet the requirements for love, and takes a form that is closer to attachment than love.

V. THEORETICAL IMPLICATIONS:

ESTRANGEMENT AND ATTACHMENT

Although there is no resoundingly agreed upon set of criteria for determining what love is, or how it ought to be, many accounts of love discuss similar variations of requirements for ideal love. Yao and Matthes both emphasize fostering connection or communion with the beloved, and doing so in part by pursuing their best interests. Yao goes further by specifying that love should not foster estrangement from the beloved. Another aspect that is emphasized by both philosophers is being attentive to the beloved's true qualities. Although these criteria don't give us specific insight into the actual *nature* of love (whether it's an attitude, feeling, relationship, etc.), they do highlight certain attitudes that are an important part of love. In this section, I hope to show that malicious moral flaws in the beloved actually compromise one's ability to foster

communion with them, thereby compromising an important tenet of love and raising the possibility that malicious moral flaws preclude us from connecting to the beloved in a way that “true love” requires. Both Yao and Matthes attempt to give an account of love that is compatible with serious moral failure, but I believe they are mistaken in doing so. Malicious moral failure is incompatible with genuine love. But obviously, we still need some explanation of what undergirds close relationships that involve malicious moral flaws, if it is not genuine, mutual love.¹³ I will argue that attachment is sometimes a more appropriate term to use for the relationship between a lover and a malicious beloved, especially in cases of abuse.

As I’ve specified previously, malicious moral flaws in a relationship can refer to behaviors that are directed towards the relationship participant, and as such, explicitly affect them. Relationships that are characterized by this sort of flaw are often abusive. But malicious moral flaws can also refer to behavior that is directed at people outside of the relationship and that does not directly affect the relationship participant whatsoever. I believe that malicious moral flaws cause estrangement regardless of where the resulting behavior is directed, unless both parties are engaging in moral failure. I will first address malicious moral flaws that result in behavior directed at one’s relationship participant.

To address the first part of my argument, I will refer to an altered version of the example from Robinson’s *Home* that Yao uses for her project. In the original example from the text, Jack and Glory are able to cultivate a close relationship, and their familial love fosters connection

¹³ I use the phrase “genuine love” to refer to something like “healthy, albeit imperfect love.” I do not believe it is required for love to be perfect or ideal—love certainly never is. But I do believe a mutually loving relationship must meet certain minimal criteria. I also am open to the possibility of redemptive love, in which love is extended by an individual to an undeserving beloved, and it inspires the beloved to improve, etc. But recall that the focus of this project is *loving relationships*, not just the extension of love. I believe redemptive love generally does not qualify as a mutually loving relationship, because one of the parties is unable to robustly participate, until they are “redeemed,” so to speak.

between them despite Jack's many flaws. But a notable factor here is that Jack's more serious flaws are not directed towards Glory, and given her kind nature, they do not have an adverse impact on Glory's ability to bond with Jack. In cases of abuse, the flaws of one party in a relationship inhibit the ability of the other party to foster connection. Let's imagine a more pertinent scenario in which Jack often physically harms Glory when they are together; when he becomes frustrated with Glory, he punches or scratches her. It seems to me that in this situation, Glory's ability to foster a connection with Jack is severely harmed by his actions. When Jack abuses Glory, he compromises the safety, communication, and equality in their relationship. And because Glory and Jack are so close, Jack is in a more intimate position to exercise control over Glory.

Obviously, in this altered version, the criterion that love fosters connection and not estrangement is not being met. I would also argue that Glory fails to see Jack as he truly is, thereby failing to meet the second criterion for love. Now, one might say that it's not the love that's fostering the estrangement, but rather, Jack is fostering the estrangement and Glory's love is holding it together—so her "love" is still genuine. This is an inaccurate way to interpret the situation. The "loving" nature of the relationship itself is the breeding grounds for the abuse that leads to estrangement. Glory's love and care for Jack, among other factors, prompts her to stay close to him and continue to endure harm. In this way, Glory's "love" for Jack fosters estrangement as it causes her to perceive Jack in an unduly positive light. By continuing to extend love, her view of Jack is inevitably distorted favorably.¹⁴ She fails to see him as he truly

¹⁴ Perceiving a loved one as better than they truly are is common and, in many cases, harmless, and even sweet. This example differs notably in that Glory's love for Jack heavily distorts her view of him, leading to a positive assessment of his character when it ought to be seen as negative and dangerous.

is—a malicious person.¹⁵ Seeing someone for who they truly are is not merely being able to recite the correct behaviors, rather, it involves fairly and accurately assessing their character. It seems that Glory does not truly see Jack as he is (a dangerous, vicious/malicious person). If she did see Jack for who he truly was, she would likely disengage from their loving relationship. This does not mean that she would cease to care about Jack, or feel attachment towards him. But if she fully understood the severity of his violation of her trust and well-being, I believe she would also see the need to disengage from Jack. In sum, due to Jack’s behavior, Glory is prohibited from having a safe, close relationship and is thereby estranged from Jack. Additionally, her love for Jack prompts her to be overly charitable to Jack, and to fail to see him for the vicious, dangerous person he actually is.

I will now elaborate on estrangement in cases where the malicious behavior is not directed at the relationship participant. To help illustrate this point, I will turn to a different case. I will draw my example from Anna and Henry’s relationship in the classic film *The Third Man* as it appears in Julia Driver’s essay “Love and Duty.” In *The Third Man*, Anna discovers that Henry, the man she loves, is responsible for the deaths of many children that were treated with the diluted black-market penicillin that he provided (Driver 6). Anna was not directly harmed by Henry knowingly profiting off of the dangerous medication he provided the children. But it still seems *very* relevant to the status of their relationship; it would be quite odd for Anna to claim that Henry’s actions don’t matter to her whatsoever. And yet, that is just what Anna does. Anna continues to love Henry, and even aids him in avoiding justice out of love for him (Driver 6). By

¹⁵ I acknowledge that to sum an entire person up as “malicious” is overly simplistic. Humans are complex, and even deeply, maliciously flawed people may have some redeeming qualities. Given the severity of the case I’m discussing here, I believe it is fair to say that Jack is a relevantly malicious person and that Glory should understand him as such, if only to preserve her own safety.

doing this, Anna becomes a willing participant in the malicious wrongdoing. Yes, she appears to successfully maintain connection with Henry, but only at the cost of her own moral standing.

And just as Glory fails to truly see Jack for the malicious person he is in our previous example, Anna fails to see Henry for the malicious person he is by finding ways to justify his actions, and thereby becomes malicious in her complicitness.

With this in mind, my claim regarding malicious moral flaws not directed at relationship participants is more nuanced than my claim regarding obviously abusive relationships. When one identifies this kind of malicious moral flaw in their beloved, they inevitably take one of three courses of actions: (1) they can “turn a blind eye” to the wrongdoing and thereby estrange themselves by refusing to be complicit in the wrongdoing and not being attentive to who their beloved truly is, (2) they continue to foster attachment and attentiveness by becoming actively complicit in the wrongdoing, or (3) they disengage from the relationship.¹⁶ It seems that in Anna’s case, upon discovering Henry’s responsibility for the death of several children, she could either turn a blind eye and distance herself from Henry, which would compromise her attitude towards him as loving (as determined by Yao’s and Matthes’s criteria), or she can continue to foster connection with Henry and participate in covering up and justifying the wrongdoing alongside him. I do not see a scenario in which a person can truly be attentive and connected to a person whose egregious wrongdoing they are aware of without becoming complicit in it. Any practice of turning a blind eye or being unwilling to endorse the wrongdoing would foster at least

¹⁶ I would argue that phenomena like redemptive love, as mentioned in an earlier footnote, qualify as a combination of these actions, primarily (1) and (3). The lover must partially disengage from the beloved’s projects and interests to avoid being complicit in the wrongdoing, but they are also able to see past the wrongdoing and focus on other redeeming aspects as they extend love. Thank you to Justin White for raising this concern.

some level of estrangement. This presents the lover with a dilemma: either become a participant in the malicious moral flaw themselves or damage the relationship they have with the beloved.

Yao would likely disagree with my claim here. She might claim that merely because one chooses to love their beloved for their flaws does not mean they are complicit in wrongdoing or morally responsible for that wrongdoing. And that may be right for the more minor offenses she discusses. But the issue is that none of the examples Yao uses are severe enough to illustrate why her view is problematic. Glory being complicit in Jack's vanity or mischievousness, whether or not it's advisable, is not obviously problematic. On the other hand, Anna being complicit in Henry causing the death of multiple children seems deeply problematic, and certainly has implications for her own moral standing. The latter example can help us see more clearly the impact that moral flaws have on the status of loving relationships.

As for Matthes, we may be harder pressed to take issue with his account on the grounds that loving a malicious person fosters estrangement, primarily because he is willing to grant that often loving relationships must compromise one or the other of his proposed criteria for love; after all, the people that engage in loving relationships are imperfect and the love that results will likewise be messy and imperfect (Matthes 252). It is certainly true that no person can boast of a perfect attentiveness to their relationship partner, or a perfect connection with them. But what is the point of having criteria if a person can fundamentally fail to meet *all* of the criteria and still meet the status for which the criteria were created? I believe the cases that I have discussed in this section exemplify the beloved's failure to satisfy any of the requirements in Matthes's account. Yes, Matthes is right that we must be willing to engage in imperfectly loving relationships, but he is wrong to think that compromising important criteria is compatible with love in every instance. Once again, Matthes fails to consider the implication of more serious

moral flaws for his theory. Malicious moral flaws do not merely bend the criteria, they promote the opposite result of the criteria. It would be unreasonable to imagine that such obvious failure to meet central elements of love is permissible for a relationship to continue to be classified as loving.

So, if a relationship with a maliciously flawed person is incompatible with genuine love, what sort of relationship is it? Obviously, in these cases, there is an important connection that constitutes the relationship. But as I've shown, it is not a genuinely loving connection. I argue that in these cases, what is often called love is better classified as attachment. Love and attachment are nearly indistinguishable, phenomenologically speaking. In fact, attachment is involved in most cases of love. To love someone without being attached to them in some manner would be quite unusual (excluding Kierkegaard-style love for your neighbor out of duty). But, one can easily be attached to someone without loving them, seeing as the criteria for attachment are far easier to meet than the criteria for love. Thus, many cases of abusive relationships, or relationships involving other malicious moral flaws, are better described as attached relationships than loving relationships, because the relationship participants fail to possess important elements of love while still maintaining an attachment to one another.

Monique Wonderly makes the claim that attachment is an "attitude marked primarily by self-focused emotions and emotional predispositions" in her essay "Love and Attachment" (235). According to her, the criteria for attachment are as follows:

- I. The attached party desires engagement with a non-substitutable particular.
- II. The attached party suffers a reduced sense of security upon prolonged separation from the object or even at the prospect of such separation.

III. The attached party experiences an increased sense of security upon obtaining the desired engagement with her attachment object. (242)

As we can see, for someone to be attached to another person, they must be pained upon separation from the person, and relieved at the presence of that particular person. Attachment posits no requirements regarding how well you understand the person you're attached to, your treatment of them, or how closely connected you are to them in reality. Attachment does not necessarily suffer at the hands of estrangement or deluded perceptions of character. It is an emotional state that can be instantiated without any regard for the nature of the relationship to the object of attachment.

Another word often used to describe attachment is *caring*. I will not draw a particular distinction between caring and attachment, although I'm sure many could be drawn.¹⁷ I argue that both caring and attachment are more apt ways of describing a person's attitude towards their malicious beloved, rather than love. To see this in application, let's return to my earlier examples. We've established that, in my modified example, Glory's attitude toward Jack cannot be genuinely loving in the sense that Matthes or Yao wants, because she fails to see him as he truly is and is estranged from him because of his violent behavior towards her. But Glory certainly desires engagement with Jack, since she remains close to him despite the abuse. As such, she acts in a way that indicates she would be pained by losing Jack. It seems clear that Glory is attached to Jack, but she is unable to foster a loving relationship with him due to his actions.

¹⁷ For instance, Wonderly argues that "we care about someone or something insofar as we desire for that person or object to flourish . . . being attached to, as opposed to caring for, another is an essentially self-interest attitude" (243).

In Anna's situation, she opted to become an active participant in Jack's malicious moral flaw. In doing so, she removed herself from the scope of my claim that it is not possible to genuinely love a malicious person. Because Anna was willing to develop a malicious moral flaw of her own to maintain a loving relationship with Henry, she is able to continue loving him at the cost of her own moral standing. But my project is primarily concerned with relationships in which only one relationship participant is malicious. Of course, the realm of possibilities is different for relationships in which both parties are malicious, as becomes the case for Anna and Henry. But let's imagine that Anna, rather than aiding Henry in avoiding justice, turns a blind eye to the wrongdoing, and tries to convince herself that she never discovered it. She no longer has a fully loving attitude towards Henry since her decision to "look away" estranges her from Henry. But she is obviously still attached to Henry, shown by her continuing to engage in a romantic relationship and seek Henry's companionship. Attachment is compatible with her intentional estrangement. Love is not.

In this section, I have attempted to show how malicious moral flaws compromise one's ability to foster communion with the beloved and/or see them as they truly are, and in doing so, compromise one's ability to love malicious persons in an ideal manner. I have also referenced a version of attachment theory that better describes what occurs in abusive relationships. But illustrating the theoretically problematic nature of remaining attached to a malicious person is not sufficient for the aims of this project. Indeed, my desire for a practical solution is what brought me to take issue with accounts like Yao's and Matthes's in the first place. It is not enough to instruct that a certain "mindset" of permissivity or dismissivity be adopted towards certain moral flaws. An adequately responsive theory requires some admonition of *action* to be taken in regard to the relationship. It seems to me that there is a deeply normative element of

loving relationships that should not be disregarded. With this in mind, I aim to propose a method of response that encourages victims of abuse and others affected by malicious behavior to disengage from the relationship and thereby rescind the love they have extended to that relationship. By providing this account in what follows, I hope to propose a method that can avoid the damage that comes from being complicit in malicious behavior but also account for the difficulty of severing loving relationships.

VI. PRACTICAL IMPLICATIONS:

VULNERABILITY TO A RELATIONSHIP

At this point, the foundation for the definition of love we've been relying on may feel shaky. So far, I have adopted the notion of love found in accounts like Yao's and Matthes's that love is broadly constituted by attitudes. This formulation of love was useful for highlighting problematic theoretical implications that malice has for genuine love. Surely, attitudes can be loving, and this aspect of love is often central. But if we merely characterize love as something that one feels an inclination in regard to (be it an attitude, valuing, feeling, etc.), we accept that the way we love others is largely out of our control. I do not believe that any of these characterizations provide a complete or robust definition of all that love encapsulates. In order to better understand what love is, and how we can use this knowledge to respond to moral flaws, I will turn to Niko Kolodny's influential account of love.

In "Love as Valuing a Relationship," Niko Kolodny argues that love is rendered normatively appropriate by the presence of a relationship (146). Moreover, he says that "love partly consists in the belief that some relationship renders it appropriate, and the emotions and motivations of love are causally sustained by this belief" (146). Through this lens, he views love as a sort of valuing, namely, valuing a relationship in a certain manner. So, when we love a

person, there is not just one foci (the beloved), but two foci (the beloved and the relationship with the beloved as its own entity). Kolodny argues that to truly constitute love, both the person and the particular relationship you have with them must meet certain criteria. Thus, if *A* is in a relationship with *B*, *A* has duties to both *B* and *R*, with *R* being the relationship itself. According to Kolodny, for *A* to value its relationship to *B* so as to constitute love, *A* must (in a significantly simplified format):

- I. Believe that it has a finally valuable relationship (*R*) with *B*;
- II. Be emotionally vulnerable to *B*, and believe that *R* is a noninstrumental reason for doing so;
- III. Be emotionally vulnerable to *R*, and believe that *R* is a noninstrumental reason for doing so;
- IV. Believe that *R* is a noninstrumental reason for *A* to act in *B*'s interests; and
- V. Believe that *R* is a noninstrumental reason for *A* to act in *R*'s interests. (151)

As is apparent from these criteria, the way in which a person perceives and acts towards the relationship participant (whether romantic, familial, or otherwise) can be separated from the way they perceive and act towards the relationship itself. Interestingly, vulnerability towards both the beloved and the relationship is critical for valuing a relationship in a loving way. Kolodny clarifies what precisely he means by vulnerability by stating the following: “to say that *A* is emotionally vulnerable to *B* (or *R*) is to say, in part, that *A* is disposed to have a range of favorable emotions in response to *A*'s beliefs that *B* (or *R*) has fared or will fare well, and a range of unfavorable emotions in response to *A*'s beliefs that *B* (or *R*) has fared or will fare poorly” (152). In other words, when *B* is joyful, or succeeds at something, *A* shares in that joy because of their vested interest in *B*'s flourishing. But of course, *B* might be flourishing while *R* is decidedly

not flourishing. Hence the need for separating the two. *A* is affected by *B*'s state of wellness, and *R*'s state of wellness.

Another important aspect of love that Kolodny highlights is that *A* acting in the interest of *B* might also include "protecting or promoting what matters to *B*," even if this may be something other than, or even contrary to, *B*'s well-being (152). This distinction highlights an interesting feature of love: when we love someone, we desire the happiness and flourishing of our beloved by their own lights, even if what they see as their happiness and flourishing may not always be in their best interest. This does not mean that we are choosing to promote something contrary to our loved one's best interests because we believe the interest is good. But we are naturally affected by the emotional state and disposition of our loved ones, and we want to respect their autonomy in forming interests and projects. Think of a mother who, knowing her child enjoys playing a particular video game, allows him extra screen time because it brings her joy to see him happy, and she knows additional gaming time will make him happy. She does this despite believing that any extra screen time over her predetermined limit goes against her son's best interests. Or think of an older sister who takes an interest in listening to heavy rap music because her brother is passionate about it and she wants to connect with him, despite her believing that its vulgarity is damaging to one's mental state.

In both examples above, the mother's and older sister's reasons for action depend on their relations to their child and brother, respectively. If we were to remove those relations from the scenarios, their behaviors would no longer make sense. Kolodny thinks this is due to what he calls "agent-relativity," meaning that the mother's or older sister's reasons for action are relative to a certain agent because of their participation in a relationship with that agent (153). Agent-

relativity can also give some insight into how ceasing to engage in a relationship changes one's reasons for action and, in turn, may limit the behaviors that those reasons prompt.

With this in mind, we can return to the project at hand. In what remains, I want to establish that, with a Kolodnian picture of love, we can rescind our love from a relationship even if we are unable to rid ourselves of a loving emotional response to a certain person. I also want to establish that, in almost every case in which our beloved has malicious moral flaws, we *ought* to disengage from our relationship with them. To do so, I will draw on Julia Driver's essay "Love and Duty." The focus of Driver's project is different from my own, since she is concerned with defending the view that we have no duty to love a particular person, and her claim that it is possible to have a duty *not* love is merely instrumental in establishing her thesis that we cannot have a duty *to* love. In drawing on Driver, I hope to meaningfully add to her account of one's duty not to love.

As Driver aptly puts it, "One cannot directly, or 'at will,' control an emotional response, but one can undermine any commitment one would normally have" (2). In this way, "being disposed to act a certain way with respect to the person one has the feelings for is controllable" (2). When we are in a loving relationship with someone, we make an emotional commitment to them. In this way, we are quite vulnerable to our loved one. If we cease to engage in that relationship, we may feel an emotional attachment or inclination towards the loved one, but we have rescinded our emotional commitment, and thereby are not as vulnerable because we are no longer privy to the aims and interests of the beloved in the same way. This is why Kolodny's emphasis on relationships is so insightful. Merely using an attitude or feeling to define love fails to capture the full picture of love and inadvertently fails to give victims of abuse a "way out" of loving a dangerous, malicious person.

Ending a relationship is not incompatible with continuing to extend love, or have a loving attitude towards a person. But the lack of a mutually loving relationship, according to Kolodny, prevents robust love. In section V, our earlier criteria for love required the ability to foster communion and see the beloved as they truly are. But both of these criteria first require that a relationship exist, on which basis to connect with the beloved. If I cease to engage in a relationship with someone, I am only capable of loving them from “afar,” so to speak, through my loving attitude towards them. This situation seems vaguely like a person who “loves” a celebrity or other public figure but does not have a personal relationship with them. Yes, I understand that this is somewhat disanalogous because in the first situation the person is equipped with all of the knowledge about their previous relationship partner to justify loving them in a personal manner. But I think the comparison highlights how important relationships are to love. Without a certain type of relationship, we lose our ability to genuinely and *personally* love others.

Fortunately, our engagement in relationships is something that we have some control over, for the most part. Disengaging from a relationship may look different depending on the type of relationship. For instance, familial relationships can be more difficult to fully disengage from than friendships if other obligations require that you continue to see your family. Or in some seriously abusive relationships, the victim may be unable to disengage from the relationships for fear that their immediate safety would be at risk. Regardless, we are by and large capable of exercising control over the loving relationships we have in our life. And because personal love to some extent requires the presence of a relationship, we are capable of exercising control over our love for others, even if it’s not in the way that we often think. In Kolodny’s terms, we can come to understand that R (the relationship with B) is not finally valuable, and that

R is not a noninstrumental reason for being vulnerable. We can choose to no longer pursue *R*'s best interests, all while still feeling a pseudo-loving attachment to *B*. We can, in short, meaningfully rescind love when we stop valuing a relationship, even when we are unable to control the emotional attachment that we have to a certain person.

Now that I've established the possibility of disengaging from our loving relationships and the control it gives us over who we love, I will turn to the normative part of my claim. I argue that if we identify a malicious moral flaw in our loved one, we ought to change our engagement in the relationship we have with them. In most cases, we should completely disengage from the relationship in order to preserve our own safety (in the case of abuse or other potentially violent flaws) and moral standing (when we could become complicit in immoral behavior). In some rare cases, which I will expand on below, partial disengagement might be appropriate when the lover is able to cease being vulnerable to certain bad projects without compromising the relationship. But I do not believe that it is *ever* appropriate to maintain complete vulnerability to a relationship with a person who possesses a malicious moral flaw.

Disengaging from a relationship involves physical and emotional disengagement as a result of ceasing to communicate with or spend time with the beloved. This importantly puts distance between the lover and the beloved in a way that makes it much easier to rid oneself of the dispositions that usually accompany love. This is an important element, since having such dispositions toward maliciously flawed persons is often harmful in nature. Driver highlights the relevance of dispositions when establishing what is involved in the duty not to love: "the duty to not love does not involve a duty to eradicate one's feelings for someone, but will involve a duty to . . . subvert any dispositions that come along with those feelings—dispositions to seek the wellbeing of the loved one when that conflicts with justice, and dispositions to make excuses,

and ignore the evidence” (7). In fact, dispositions like ignoring evidence and making excuses involve an effort to *not* see the beloved as they truly are, which is antithetical to love. Subverting such dispositions is tied up with Kolodny’s vulnerability criteria, since vulnerability facilitates dispositions oriented towards the interests of the maliciously flawed beloved. By looking more closely at the dangers of vulnerability and dispositional inclinations towards a maliciously flawed person, we can better understand why we have a normative responsibility to disengage from them when possible and appropriate.

As I discussed earlier, Kolodny’s theory of love hinges on vulnerability. Being in a loving relationship opens both participants to some level of vulnerability by virtue of their closeness and the commitments that most relationships involve. This vulnerability is most easily identified in romantic relationships, but it also must hold true for any loving relationship between family or friends. Part of this vulnerability comes from making a conscious effort to be open and vulnerable towards a partner, but a large part of vulnerability to the beloved in a relationship stems from the very nature of relationships; because of the emotional and situational ties that a relationship involves, we are naturally vulnerable to the successes and failures of our beloved. We are also vulnerable to their actions, tendencies, and qualities, as many of these qualities will inform the way they behave towards us. This sort of vulnerability is a beautiful part of what constitutes loving relationships: we open ourselves to the beloved and all of their odd tendencies and cares. We partake of their joys and weep with them in their sorrows as we come to understand the way they experience the world. But this vulnerability is also what makes engaging in a relationship with a maliciously flawed person so dangerous. I believe that cases in which malicious flaws manifest towards the relationship participant and cases where they do not

are differently problematic and thus need to be addressed separately. I will first address the former.

It seems relatively intuitive to suggest that victims of abuse ought to disengage from their relationship with the abuser to the greatest extent possible. Vulnerability towards an abusive person is quite obviously harmful and dangerous. Abuse not only represents an egregious cruelty, but also a violation of the very relationship itself (Driver 6). Abuse in a loving relationship wholly compromises the love and trust once found within that relationship, and establishes an oppressive power dynamic that is inappropriate in any setting, much less a loving relationship. Because of the personally damaging and dangerous nature of abusive relationships, the victim has a responsibility to fully disengage from the relationship. Assuming that we have a basic duty to preserve our own safety and well-being barring conflicting duties, we have a duty to disengage from relationships in which abuse becomes apparent.¹⁸ In such cases, we ought to look out for our own well-being by ceasing to spend time with or communicate with the malicious person. It is only through disengaging in this way that one can remove oneself from abuse and cease to be vulnerable to the negative qualities of their loved one.

In the case of our earlier example where Jack was physically violent towards Glory, I argue that Glory ought to cut ties with Jack entirely in order to preserve her own safety. She owes it to herself to remove herself from a relationship in which she is consistently in harm's way. It is sometimes appropriate to put oneself in harm's way for altruistic ends. But even if we were to craft an outlandish counter-example in which some positive result comes from enduring abuse, typically the primary end achieved by enduring abuse is accommodating the desires of the

¹⁸ I will not defend this assumption here since it is largely out of the scope of my project. But for a robust defense of duties to the self and their limitations, see Yuliya Kanygina's essay on "Duties to Oneself and Their Alleged Incoherence."

abuser at the expense of one's safety, which is far too high a price to pay. Any attempt at justifying remaining in an abusive relationship by claiming that it's an appropriate case to put oneself in harm's way clearly fails.

I have now established that there is a responsibility to disengage from relationships in which malice is directed at the relationship participant. Showing that we ought to disengage from relationships where malice is not directed at the relationship participant is certainly a higher bar, since it is less obviously harmful. Let's first return to Driver's concern with dispositions. Driver believes that for cases where one loves a decidedly bad person, "blame is still warranted in such cases because love reveals one's evaluative commitments in a way that reveals something about a person's dispositions and character" (7). She aptly uses Henry and Anna's relationship to illustrate her claim, since Anna's love for Henry reveals her willingness to conceal evidence of Henry's misdeeds and thereby to participate in committing serious misdeeds of her own. Upon discovering Henry's diluted penicillin scheme and the tragedies it caused, Anna should have disengaged from her relationship with Henry. Even apart from her going on to conceal his wrongdoing, it seems wrong for Anna to love Henry. Of course, when I say it is wrong for Anna to love Henry, I am assuming a Kolodnyan definition of love. So, in other words, my claim is that it is wrong for Anna to remain in a loving relationship with Henry. The former and latter claim are synonymous according to Kolodny, but it is still an important clarification to make.

I argue that the wrongness of Anna loving Henry is grounded by the morally fraught dispositions and vulnerability to bad aims that loving a maliciously flawed person facilitates. It is very difficult to entirely maintain one's own moral standing while fostering connection in a loving relationship with a malicious person. This is because fostering that connection requires vulnerability, and vulnerability involves being affected by our beloved's interests. When we are

affected in this way, we work to promote and protect those interests. This even applies to interests that are not actually in the best interest of the beloved. Being affected by such interests is relatively innocent when it involves allowing more screen time or occasionally listening to vulgar music. But when we are vulnerable to the aims and interests of a malicious person, the resulting affective dispositions are not so innocent. Anna conceals evidence of Henry's crime because she wants Henry to be comfortable, safe, and happy; Henry being in a state of well-being makes Anna happy because she loves him. Her being affected by Henry's happiness in this way is normal, but the resulting behavior is *not* normal, it is morally reprehensible. Love is tricky; we cannot choose "at will" to be unaffected by our beloved's interests simply because they are malicious. When we open ourselves to an emotionally vulnerable relationship, we may end up caring about all sorts of things, good and bad. Since we ought to avoid being affected by the bad aims and interests of a malicious person, we ought to disengage from our relationship with them.

A possible objection to this argument is that Kolodny should not be interpreted to claim that we are vulnerable to interests of the beloved that are against their *actual* best interests. An alternative reading of Kolodny might suggest that his opting for the phrase "interests" rather than "*best* interests" is simply meant to allow for neutral or minorly harmful interests. One may claim, as Matthes does, that since most everyone has an interest in flourishing, and immorality impedes flourishing, immoral behavior is never *really* in anyone's interest (Matthes 246). Perhaps Henry isn't really happy; he would only experience true happiness if he repaired his wrongdoing. To me, the idea that we can know the true best interests of our relationship partner is somewhat presumptuous. Sure, being in a relationship with someone does give a person unique insight into the interests and needs of the beloved, but it does not give the lover unequivocal access to what

will make their partner “truly” happy. The claim that the relationship participant with the moral high ground always knows what’s really best leads to an odd breed of paternalism. Of course, it would be best if Henry had not sold the diluted penicillin. But what is best by Henry’s own lights is obviously still relevant. It does not hinge any less on personal feelings than what is best by a morally upstanding person’s lights. Further, I reject the notion that for *every* moral agent, truly being happy involves living a moral life. It is very possible that Henry would not be happier living a morally upstanding life, even if he ought to be. It is not clear to me how we can exclude self-perceived interests and happiness from affecting the dispositions of a malicious person’s relationship partner. Perhaps Anna has a different idea of what would truly make Henry happy, but that does not prevent her from being affected by what Henry has decided makes him happy. And Anna does end up being affected by Henry’s interests: she protects his bad interests and covers for his crimes because her own happiness hinges on Henry’s contentment. This results in her harboring decidedly bad dispositions (e.g., willingness to circumvent justice, etc.). Vulnerability to the beloved’s interests and the dispositions that result from that vulnerability are important and often inevitable parts of loving relationships.

Another possible objection one might make is that Henry’s wrongdoing is simply not relevant to his relationship with Anna, and therefore Anna is not obligated to respond to the wrongdoing. Perhaps Henry has other great qualities that are more relevant to his behavior in personal relationships. Driver articulates this objection in the following terms:

It may also be reasonable for [Anna] to continue to love [Henry] even upon discovering he is a murderer, since there may be other qualities she is responding to that have not changed (his sense of humor, etc.) . . . we also have an account of why it would be reasonable for Anna to fall out of love with Harry: there are certain relationship defining

demands that Harry has failed to live up to. One might argue that this is mistaken, though, since Harry's killing is not something that figures into his relationship with Anna. (7)

Anna and Henry's relationship seems like a pretty poor example to use if someone wanted to defend the objection Driver raises, since Henry's indirect murder of the children figures prominently in his relationship with Anna as she is involved in covering up the crime. But I can imagine a stronger example in which the wrongdoing is genuinely absent from the relationship despite being known to both parties. Let's say that Anna knew that Henry killed several children by selling his penicillin, but had decided to adopt the mindset of "what happens at work stays at work" and stopped asking questions about Henry's medical malpractice. Would she really be able to remain uncorrupted by Henry's wrongdoings? It's possible, but certainly not advisable to attempt.

I mentioned earlier that I allow for some exceptions to the notion that we ought to fully disengage from relationships with maliciously flawed persons. In select cases, partial disengagement may be appropriate. Perhaps Anna hopes to continue loving Henry, and in loving him, inspire him to cease his shady business practices. I am sensitive to the idea that forgiveness and love are important parts of redemption from serious moral flaws. I want to clarify that the aim of my project is not to claim that the solution to moral failure in loving relationships is "Don't love, shun!"¹⁹ For one, all of my normative claims in this section are limited to extremely serious cases of moral failure, which is why I have exclusively discussed malicious moral flaws. Further, to say that disengaging from a relationship must involve shunning a person, or any other

¹⁹ Credit to Derek Haderlie for raising this concern and the accompanying clever phrase.

negative attitude is simply incorrect. One need not despise or disdain someone to understand that they ought to disengage from a relationship with them.

Although this is the case, I do believe that it is advisable to avoid corrupting or endangering oneself by being in a relationship with a malicious person rather than remaining vulnerable to them under the assumption that you can “fix” the malicious person by extending love to them. There are select cases in which it would be appropriate to stay partially engaged with a malicious person to extend forgiveness and love in an attempt to encourage moral improvement. If Anna is genuinely unaffected by Henry’s wrongdoing, she may be in a position to see his redeeming qualities and love him without justifying or condoning his flaws. She may be able to avoid being vulnerable to Henry’s happiness or success as it pertains to his bad aims or interests, while still being vulnerable to his good aims and interests. As I emphasized previously, consciously determining which aspects of a relationship and/or person one is vulnerable to is *extremely* difficult. But in an effort to allow for redemptive love, I will allow for the possibility that one could love a maliciously flawed person from a place of partial disengagement without themselves being compromised.

In this final section of my thesis, I have attempted to defend the normative element of my theory for responding to malicious moral flaws in loving relationships. Adopting Kolodny’s view of love as valuing a relationship allows for victims of abuse to cease loving malicious persons, but most importantly, it highlights why we *ought* to disengage from relationships with maliciously flawed persons. We have a responsibility to stop being vulnerable to maliciously flawed people. We owe it to ourselves to protect our own safety and avoid immoral dispositions. I believe that this normative element is essential to the success of my project. Much of my project is descriptive in nature: defining different types of moral flaws, highlighting the ways

that malicious moral flaws compromise genuine love, etc. These descriptive elements all work together to provide the foundation for my central claim that there are genuine practical implications of engaging in loving relationships with maliciously flawed people.

VII. CONCLUSION

I do not claim to have offered the solution for all, or even most of the difficulties that malice causes in loving relationships, or to have explored every way in which love is altered by abuse. I *have* attempted to provide a careful groundwork for a theory of love that can respond to serious moral failure, both descriptively and normatively. Yao and Matthes highlight central elements of love and how it manifests in interpersonal relationships. They speak of attentiveness, grace, concern for well-being, and being accepting of imperfection. Without these features, we would surely have an impoverished understanding of love. But while many of these features are necessary, they are not sufficient for an account of love that is sensitive to malicious moral flaws that result in abuse and other immoral behaviors.

Indeed, this thesis has focused on the less palatable cases that lurk on the outskirts of love. While it may seem that I have dedicated an undue amount of space to discussing flaws that fall quite far on the extreme side of the spectrum of moral failure, their presence on the extreme side does not mean they are not prevalent in loving relationships of all varieties. Not only do malicious moral flaws in loving relationships compromise our ability to love by estranging us from the beloved, they also imbue us with a normative responsibility to disengage from our relationship with the beloved given the dangers of remaining vulnerable. By disengaging in this way, we can effectively stop loving the malicious person without setting an unrealistic imperative to extinguish loving feelings or attachment towards our malicious loved ones. I hope that I have adequately illustrated why these arguments hold true not only in cases of abuse, but

also in cases where malice manifests outside of the relationship. It is only when we expand our notion of moral failure in loving relationships to encapsulate these cases that we can conceptualize the sensible limits of love.

WORKS CITED

- Driver, Julia. "Love and Duty." *Philosophic Exchange*, vol. 44, no. 1, 2014.
- Kolodny, Niko. "Love as Valuing a Relationship." *Philosophical Review*, vol. 112, no. 2, 2003, pp. 135-189.
- Korsgaard, Christine M. *The Sources of Normativity*. Edited by Onora O'Neill, Cambridge University Press, 1996.
- Matthes, Erich Hatala. "Love in Spite of." *Oxford Studies in Normative Ethics*, vol. 6, 2016, pp. 241-262.
- Nietzsche, Friedrich. *On the Genealogy of Morals*. Translated by Douglas Smith, Oxford University Press, 1996.
- Tolstoy, Leo. *Anna Karenina*. Translated by Richard Pevear and Larissa Volokhonsky, Penguin Classics, 2000.
- Wonderly, Monique. "Love and Attachment." *American Philosophical Quarterly*, vol. 54, no. 3, 2017, pp. 232-250.
- Yao, Vida. "Grace and Alienation." *Philosophers' Imprint*, vol. 20, no. 16, 2020, pp. 1-18.